

# CONTINUITY OF THE LYAPUNOV EXPONENTS OF RANDOM MATRIX PRODUCTS

ARTUR AVILA, ALEX ESKIN, AND MARCELO VIANA

ABSTRACT. We prove that the Lyapunov exponents of random products in a (real or complex) matrix group depends continuously on the matrix coefficients and probability weights. More generally, the Lyapunov exponents of the random product defined by any compactly supported probability distribution on  $GL(d)$  vary continuously with the distribution, in a natural topology corresponding to weak\*-closeness of the distributions and Hausdorff-closeness of their supports.

## CONTENTS

<b>Part I. Lyapunov exponents and random walks</b>	2
1. Introduction	2
Lyapunov exponents	2
Continuity theorem	3
Margulis functions	4
Further perspectives	5
Significance and applications of continuity	6
Quantitative regularity	8
2. Statement of main result	9
3. Invariant subspaces	11
3.1. Stationary measures	12
3.2. The equator	13
4. Uniform convergence in measure	15
<b>Part II. Preliminaries and outline of the proof</b>	20
5. The equator is a repeller	20
5.1. Generic measures	21
5.2. Repelling behavior	22
6. A toolbox	26
6.1. Couplings	26
6.2. Markov operators	30
6.3. Invariant couplings	34
6.4. Margulis functions	35
6.5. Adapted operators	37
7. Outline of the proof	38

---

*Date:* May 16, 2023.

M.V. was partially supported by CNPq, FAPERJ, and Fondation Louis D–Institut de France. A.E. was partially supported by NSF grants DMS-1201422, DMS-1500702, DMS-1800646 and the Simons Foundation.

7.1. Main inductive statement	39
7.2. First step	40
7.3. General step	43
<b>Part III. First step of the induction</b>	45
8. Preparing a Margulis function	45
8.1. Vertical angle function	45
8.2. Vertical projection function	47
8.3. The function $-\log \psi_1$	51
9. Stabilization and cut-off	52
9.1. Stabilization	53
9.2. Cutoff	54
10. Turning the perturbation on	57
11. Recoupling and conclusion	64
11.1. Recoupling	64
11.2. Contradicting $\dim E = 1$	67
11.3. Completing the first step	69
<b>Part IV. General step of the induction</b>	72
12. Preparing a Margulis function	72
12.1. Vertical angle function	72
12.2. Vertical projection function	74
12.3. The function $-\log \psi_r$	77
13. Stabilization and cut-off	80
13.1. Stabilization	80
13.2. Cutoff	83
14. Turning the perturbation on	85
15. Spreading out	93
15.1. Homogeneous measures on flag varieties	93
15.2. Spreading out operators	94
16. Recoupling and conclusion	99
16.1. Markov operators on flag varieties	100
16.2. Recoupling	100
16.3. Contradicting $\dim E = r$	104
16.4. Completing step $r$	105
References	109

## Part I. Lyapunov exponents and random walks

### 1. INTRODUCTION

**Lyapunov exponents.** The notion of Lyapunov exponents is rooted in the stability theory of differential equations created by Lyapunov [52] at the end of the 19th century. Consider a differential equation

$$(1.1) \quad x' = L(t)x + R(t, x),$$

where  $L(t) : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is linear and  $R(t, \cdot)$  is a perturbation of order greater than 1. The *Lyapunov exponent function*  $v \mapsto \lambda(v)$  is defined by

$$(1.2) \quad \lambda(v) = \limsup_{t \rightarrow \infty} \frac{1}{t} \log \|x_v(t)\|$$

where  $x_v$  is the solution of the linear equation  $x' = L(t)x$  with initial condition  $v$ . When  $\lambda < 0$  the constant solution  $x_0(t) \equiv 0$  is exponentially stable for this linear equation. The stability theorem of Lyapunov asserts that, under a technical *Lyapunov regularity* condition, it remains exponentially stable for (1.1).

In 1960, Furstenberg, Kesten [41] proved that the limit in (1.2) exists for almost every  $x$ , relative to any probability measure invariant under the flow. A few years later, Oseledec [56] showed that Lyapunov regularity also holds for almost every point. Such results brought the subject of Lyapunov exponents to the realm of ergodic theory, where it has prospered since. Three main problems have a central role in the theory.

The first one is *non-triviality* of the Lyapunov spectrum: *when is it the case that not all Lyapunov exponents are equal?* This was founded by Furstenberg [40] in the 1960s and has been much studied since, especially in the last couple of decades or so. See Viana [65] and references therein. A product of this theory much exploited recently is the Invariance Principle [51, 20, 7, 5], a general statement to the effect that systems with trivial Lyapunov spectra are very rigid.

A related issue is that of *simplicity* of the Lyapunov spectrum: *when are all the Lyapunov exponents distinct, with multiplicity 1?* This was initiated by Guivarc'h and Raugi [46] and by Gol'dsheid and Margulis [45], and has also been the object of considerable interest in recent years. See [65] for references and a detailed discussion. An application was the proof of the Zorich–Kontsevich conjecture on the Lyapunov spectrum of the Teichmüller flow on the moduli space of Abelian differentials [6].

**Continuity theorem.** In the present paper we are mostly concerned with the *dependence* problem: *how do the Lyapunov exponents depend on their underlying system?* Several references to the literature on this problem will be given in a while. Right now, the following special case of our main result illustrates the kind of goals we pursue here.

Let  $(A_1, \dots, A_m)$  be an  $m$ -uple of matrices in  $G = GL(\mathbb{R}^d)$  and  $(p_1, \dots, p_m)$  be an element of the open simplex  $\Delta_m$  of dimension  $m - 1$ , that is, an  $m$ -uple of numbers  $p_i \in (0, 1)$  such that  $\sum_{i=1}^m p_i = 1$ . Let  $\nu$  be the probability measure on  $G$  given by

$$\nu = \sum_{i=1}^m p_i \delta_{A_i}$$

where  $\delta_A$  denotes the Dirac mass at any  $A \in G$ . Let  $\lambda_1 \geq \dots \geq \lambda_d$  be the Lyapunov exponents of the random matrix product induced by  $\nu$  (definitions will appear in Section 2). We prove:

**Theorem A.** *For each  $1 \leq j \leq d$ , the number  $\lambda_j$  depends continuously on the  $A_i$  and the  $p_i$  at every point of the domain  $G^m \times \Delta_m$ .*

The 2-dimensional case of Theorem A has been proved by Bocker and Viana [18]. A different proof of that case that introduces a few of the ideas in this paper appeared in Chapter 10 of the book [65]. Kifer [50] observed the Lyapunov exponents

may actually jump when some weight  $p_i$  goes to zero, and that is why  $\Delta_m$  is taken to be an *open* simplex.

A crucial point in Theorem A is that the conclusion holds even at *reducible* points, that is, when the matrices  $A_1, \dots, A_n$  share one or more invariant proper subspaces. Indeed, continuity of the Lyapunov exponents at irreducible points had already been proved in the 1980s, independently, by Furstenberg and Kifer [42] and by Hennion [47]. As often happens in this field, the reducible case is a lot more subtle, requiring a whole different set of ideas.

Our approach relies on a quantitative analysis of the random walk

$$x \mapsto gx, \quad g \in G \text{ a random variable with distribution } \nu,$$

defined on the projective space  $P = \mathbb{P}(\mathbb{R}^d)$  by the probability measure  $\nu$ . By Furstenberg and Kifer [42], discontinuity of the Lyapunov exponents can only occur if there is some proper subspace invariant under all the matrices and containing all the “most contracting” directions for the cocycle (see Section 3 for the precise statement).

In a nutshell, we prove that if such an invariant subspace  $E$  (the “equator”) does exist, typical trajectories of nearby generic random walks spend very little time in its vicinity, rendering the presence of the equator effectively harmless. A bit more precisely, we consider generic distributions  $\nu_k$  converging to  $\nu$  as  $k \rightarrow \infty$ , and we show that the stationary measures  $\eta_k$  for the corresponding random walks cannot accumulate on the equator: any limit point  $\eta_\infty$  as  $k \rightarrow \infty$  (which is automatically a stationary measure for  $\nu$ ) must satisfy  $\eta_\infty(E) = 0$ . These notions and their basic properties will also be recalled in Section 3.

**Margulis functions.** The key technical tool to do this is the concept of Margulis function. Such functions have been introduced to the dynamics literature by Margulis in [32]. (In the probability setting, a Margulis function is also called a Foster-Lyapunov (or drift) function, and has been used extensively. See the book [55] for further references.)

In a few words,  $\Phi : X \rightarrow [0, \infty]$  is a (*multiplicative*) *Margulis function* for a Markov operator  $\mathcal{T}$  on some space  $X$ , relative to a set  $Y \subset X$ , if  $\Phi \equiv \infty$  on  $Y$  and there exist constants  $c < 1$  and  $b < \infty$  satisfying

$$(1.3) \quad \mathcal{T}\Phi(x) \leq c\Phi(x) + b \text{ for all } x \in X.$$

The distinctive feature implied by this inequality is that  $\mathcal{T}\Phi(x)$  is much smaller than  $\Phi(x)$  near  $Y$ , even if it may be somewhat bigger on other parts of  $X$ . Such functions have been used, for example, in [34, 31, 33, 1, 2, 24, 35, 14, 30, 3, 49, 53, 12, 43, 44] and [65, Chapter 10]. For a fairly recent survey on this topic see [38].

Nevertheless, our application of Margulis functions in the present setting comes with a number of novelties. To begin with, we need a different kind of Margulis function, which we introduce here: given a partition  $(A, B)$  of the space  $X$ ,  $\Phi$  is an (*additive*) *Margulis function* if there exist positive constants  $\kappa_A$  and  $\kappa_B$  such that

$$(1.4) \quad \begin{aligned} \mathcal{T}\Psi(x) &\leq \Psi(x) - \kappa_A \text{ for every } x \in A \\ \mathcal{T}\Psi(x) &\leq \Psi(x) + \kappa_B \text{ for every } x \in B. \end{aligned}$$

If  $\Phi$  is a multiplicative Margulis function then  $\log \Phi$  is an additive Margulis function relative to a suitable partition  $(A, B)$  (see Remark 6.22). On the other hand, it is not true that if  $\Psi$  is an additive Margulis function then  $\exp \Psi$  is a multiplicative one. That is because the inequality (1.3) is very sensitive to the

“worst case” behavior of  $\Psi$ , whereas (1.4) depend more on the “average case” behavior. For this reason, it is often much easier to construct an additive Margulis function than a multiplicative one. In fact, we do not know how to construct a useful multiplicative Margulis function in our setting beyond the case  $d = 2$  (see [65, Chapter 10] and [53]).

Another point worth emphasizing is that in all the previous constructions (apart from [53, 65, 12]), the dynamical system is fixed. Instead, in the present paper we try to make the same Margulis function work for a whole family of dynamical systems, which have very different behaviors near the equator. This introduces quite a lot of new issues. Furthermore, the dynamical behavior near the equator is totally different from the behavior in other parts of phase-space. Thus, we need to carry out a localized analysis of the random walk, which is another important source of difficulties.

**Further perspectives.** Random products of matrices may be represented as a special kind of *linear cocycle*

$$F : M \times \mathbb{R}^d \rightarrow M \times \mathbb{R}^d, \quad F(x, v) = (f(x), A(x)v)$$

where the base dynamics  $f : M \rightarrow M$  is a shift map  $f((x_n)_n) = (x_{n+1})_n$  endowed with a Bernoulli measure, and the cocycle function  $A$  depends only on the coordinate  $x_0$ . The dependence problem extends naturally to this general setting of linear cocycles: usually one takes the base dynamics  $f$  and the corresponding invariant probability measure  $\mu$  to be fixed, and one is interested in understanding how the Lyapunov exponents depend on  $A$ .

A natural step is to try and allow for much more general cocycle functions  $A$ . For reasons that we will soon discuss, it is convenient to assume some regularity, like Hölder continuity. Moreover, essentially all known results assume the cocycle to satisfy a kind of quasi-conformality condition called *fiber-bunching* (see [20, 7]) which also involves Hölder continuity.

Another natural way to broaden the scope of the theory is to weaken the assumptions on the base dynamics, to consider general dynamical systems more general than shift maps, as well as invariant measures satisfying much milder independence conditions. In this latter direction, Theorem A has been extended to Markov products of 2-dimensional matrices by Malheiro and Viana [53].

For Hölder cocycles, Backes, Brown and Butler [12] extended the 2-dimensional case of Theorem A to general fiber-bunched cocycles whose base transformation  $f$  is a hyperbolic homeomorphism on a compact metric space (in the sense of [64]) and whose invariant probability measure has *local product structure* (see [20, 64]), a mild requirement meaning, roughly speaking, that the future depends only weakly on the past. In fact, their statement extends to the class of linear cocycles with invariant holonomies. Both versions had been conjectured in [65, Section 10.6].

Another interesting path to possibly generate further progress is to consider linear cocycles over partially hyperbolic diffeomorphisms, volume-preserving or not. Groundwork in this direction has been laid in [5] and some continuity results have been derived in [7]. See also Avila, Viana and Wilkinson [8, 9], and Poletti and Viana [59].

The need for some regularity of the cocycle function is highlighted by the following result of Bochi [15]: if the system  $(f, \mu)$  is aperiodic,  $d = 2$ , and the linear cocycle

is continuous and not uniformly hyperbolic, then it may be  $C^0$ -approximated by linear cocycles with trivial Lyapunov spectra (the two Lyapunov exponents are equal). Thus, continuity can only hold at cocycles which either are uniformly hyperbolic or have trivial spectra.

In fact, the same is true restricted to the class of derivative cocycles of area-preserving surface diffeomorphisms, a much harder fact which was discovered by Mañé [54] and whose proof was completed by Bochi [15]. These results have been extended to arbitrary dimensions by Bochi and Viana [17, 16]. They are generally not true for cocycles over non-invertible maps, even in the  $\mathrm{SL}(\mathbb{R}^2)$  case, according to Viana and Yang [66].

On the other hand, the actual relevance of the fiber-bunching condition in this context is presently not entirely clear, indeed this remains one of the outstanding open questions in this area. Examples of discontinuity of the Lyapunov exponents for Hölder continuous linear cocycles which are not fiber-bunched have been found in [65, Section 9.3] and Butler [25].

We have restricted our attention to matrix groups, for good reason. While the basic concepts discussed here (such as Lyapunov exponents, Oseledets regularity, etc.) extend to the more general setting of all (not necessarily invertible) matrices, there is no hope to obtain any general regularity result for Lyapunov exponents in this more general setting.

To explain why, let us consider the Lyapunov exponents of the random product of two real matrices  $A_1$  and  $A_2$ , with probability weights  $p_1 = p_2 = 1/2$ . The Lyapunov exponents are well defined, but the bottom one is equal to  $-\infty$  if one of the matrices is not invertible. Moreover, if some finite matrix product involving  $A_1$  and  $A_2$  is zero then the top Lyapunov exponent is  $-\infty$  as well.

For instance, let  $L(\theta)$  denote the top Lyapunov exponent for

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } A_2 = \begin{pmatrix} \cos 2\pi\theta & -\sin 2\pi\theta \\ \sin 2\pi\theta & \cos 2\pi\theta \end{pmatrix},$$

viewed as a function of  $\theta \in \mathbb{R}$ . For  $\theta = p/4q$  with  $p$  odd and  $q$  a non-zero integer, we have that  $A_1 A_2^q A_1 = 0$ , and so  $L(\theta) = -\infty$ . By an upper semi-continuity argument, it follows that  $L(\theta) = -\infty$  for Baire-generic  $\theta$ . In fact, it is not difficult to give an explicit generic quantitative condition ensuring that  $L(\theta) = -\infty$ , and even a sharp one, using the easily checked formula

$$L(\theta) = \sum_{k=0}^{\infty} 2^{-k-2} \log |\cos 2\pi k\theta|,$$

which also shows that the Lyapunov exponent  $L(\theta)$  is finite (and discontinuous) at a full Lebesgue measure set of  $\theta$ .

What happens in this sort of situation is that, while one can still analyze the Lyapunov exponents of cocycles such as this one using a stationary measure  $\eta$  on the projective space, just as we do in the present paper for the invertible case (but taking care of issues such as indeterminacy), in the present setting the measure  $\eta$  becomes atomic, being the sum of Dirac masses with weights  $2^{-k-1}$  on the lines through  $(\cos 2\pi k\theta, \sin 2\pi k\theta)$ .

**Significance and applications of continuity.** Knowledge that the Lyapunov exponents are continuous at some  $f$  can in itself give information about the dynamics of  $f$ , as pointed out in Bochi, Viana [17], and abstract facts about the

existence of many (in the Baire sense) continuity points can be leveraged to a fine understanding of the dynamics from the generic point of view.

For instance, continuity of Lyapunov exponents ensures that the Oseledets decomposition varies continuously in a suitable sense: see Backes, Poletti [13]. Moreover, the convergence in the Oseledets theorem is locally uniform on the cocycle. This sort of uniformity is useful in various situations in dynamical systems. An example is the following relevant question in the ergodic theory of volume-preserving diffeomorphisms  $f : M \rightarrow M$ . See [39] or [60] for background.

Let  $K \subset M$  be a Pesin block, that is, a compact (non-invariant) set where the Lyapunov exponents are all bounded away from zero and the estimates in the Oseledets theorem hold uniformly. Over such a set, the Pesin stable and unstable manifolds are well defined and depend continuously on the point. In particular they have a definite size, and so nearby points in the Pesin block must belong to the same ergodic component. One may thus ask about the stability of Pesin blocks: is it the case that a smooth perturbation of  $g$  must also possess a Pesin block  $K_g$  nearby (in the sense that the symmetric difference  $K \Delta K_g$  has small measure) such that its Pesin manifolds are close to the unperturbed ones? This can be shown to follow from a suitable control of the dependence of the Lyapunov exponents, and in particular if the averaged Lyapunov exponents depend continuously on the diffeomorphism at the point  $f$ .

One setting where knowledge about the continuity of the Lyapunov exponents has been used as an essential ingredient in the understanding of the dynamics is in the study of quasiperiodic Schrödinger operators. Here  $f = f_\alpha : x \mapsto x + \alpha$  is a translation on a finite dimensional torus  $\mathbb{T}^d$ , and the cocycle function takes the form

$$A(x) = \begin{pmatrix} E - v(x) & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SL}(2, \mathbb{R}).$$

In this case the continuity of Lyapunov exponents with respect to both  $\alpha$  and  $A$  has been proved for analytic  $A$  and totally irrational  $\alpha$  (meaning that  $f_\alpha$  is minimal), by Bourgain and Jitomirskaya [23] when  $d = 1$  and by Bourgain [21] in the general case.

It was used, for instance, in the solution by Avila and Jitomirskaya [4] of the Ten Martini Problem, which asked whether the Almost Mathieu Operator has a Cantor spectrum (this can be rephrased as density of uniform hyperbolicity within certain one-parameter families of cocycles). It also appears prominently in the proof of the quantization of the acceleration for  $d = 1$ , which is the starting point of the so-called global theory of one-frequency Schrödinger operators (Avila [10]).

Continuity is very subtle in this context: for instance, the aforementioned result of Bourgain and Jitomirskaya does not hold when  $A$  is merely  $C^\infty$  (a result of Wang-You [67]). Continuity as a function of  $\alpha$  also may fail at rational  $\alpha$  even when  $A$  is analytic, see [23] for a discussion.

Another setting which connects with the ideas discussed in this paper is that of  $\mathrm{SL}(2, \mathbb{R})$ -actions on moduli spaces of Abelian or quadratic differentials. Let  $\eta$  be an  $\mathrm{SL}(2, \mathbb{R})$ -invariant probability measure. By Eskin-Mirzakhani [36],  $\eta$  is equivalent to Lebesgue measure on some submanifold. The Kontsevich-Zorich cocycle over the Teichmüller flow, which plays a fundamental role in the ergodic theory of translation surfaces, has a nice behavior with respect to  $\eta$ : it is basically a random matrix product (involving countably many matrices), except that the products are merely

“quasi-independent”. See [6] for an application of this idea to the issue of simplicity of the Lyapunov spectrum.

Let  $(\nu_k)_k$  be a sequence of  $\mathrm{SL}(2, \mathbb{R})$ -invariant probability measures converging in the weak\* sense to some probability measure  $\nu$ . It would be tempting to use the techniques of our paper to address the issue of continuity of the Lyapunov exponents in this context. However, it turns out that the difficulties addressed in our paper do not show up, and hence simpler techniques can be applied, as was done in Bonatti, Eskin, and Wilkinson [19]. Indeed, by Theorem 2.3 in Eskin, Mirzakhani, and Mohammadi [37] (see also Theorem 2.6 in [19]), the support of  $\nu_k$  is contained in  $\mathrm{supp} \nu$  for every large  $k$ . So, in terms of the random matrix models, if there is an invariant subspace for  $\eta$  then it is also invariant for the  $\eta_k$  for large  $k$ . This allows one to quotient out bad invariant spaces, and establish continuity by the usual Furstenberg–Kifer argument [42].

**Quantitative regularity.** Another natural question is how much can the regularity of Lyapunov be upgraded from mere continuity. An old result of Ruelle [61] asserts that if all the matrix coefficients are positive then the largest Lyapunov exponent is a real-analytic function of those coefficients. For locally constant cocycles over Markov shifts, Peres [58] has shown that if the Lyapunov exponents are simple then they depend real-analytically on the transition data, assuming the cocycle function itself is fixed.

For parametrized random matrix products satisfying strong irreducibility and the contraction property, Le Page [57] has proved that the largest Lyapunov exponent is a Hölder continuous function of the parameter. This function is even  $C^\infty$  if the probability distributions are absolutely continuous. In the opposite direction, a construction of Halperin (see Simon and Taylor [62, Appendix 3]) shows that for every  $\alpha > 0$  one can find random Schrödinger cocycles near which the Lyapunov exponents fail to be  $\alpha$ -Hölder continuous.

These results have been sharpened by Duarte and Klein, who developed a unified approach to proving generic moduli of continuity of the Lyapunov exponents for different classes of linear cocycles, both random and quasi-periodic, especially in the 2-dimensional case. See [27, 28] and also [26] for an account of their approach and many applications.

Still in the 2-dimensional case, Tal and Viana [63] have shown that Hölder continuity holds at every point where the Lyapunov spectrum is simple. This is an application of estimates obtained from the methods we develop here, namely a uniform bound

$$\eta(E(r)) \leq Cr^\beta$$

for the weight of the neighborhood  $E(r)$  of the equator relative to stationary measures  $\eta$  of nearby random walks. Tal and Viana [63] have also shown that, while Hölder continuity may fail when the two Lyapunov exponents coincide, a weaker log-Hölder modulus of continuity does hold at every point. It would be interesting to extend these results to arbitrary dimension.

Still regarding products of finitely many matrices, one problem that has proved to be very resistant to all techniques so far is whether the dependence of the Lyapunov exponent can be much better than Hölder in some non-trivial region of the parameter space. For instance, let us consider random matrix products of two  $\mathrm{SL}(2, \mathbb{R})$ -matrices  $A_1$  and  $A_2$ , with probability weights  $p_1 = p_2 = 1/2$ . Over the open set  $\mathrm{UH}$  of uniformly hyperbolic pairs  $(A_1, A_2)$ , the top Lyapunov exponent is



clearly a real analytic function. Fix  $1 \leq k < \infty$ . *Is there a non-empty open subset in the complement of UH over which the top Lyapunov exponent is  $C^k$ ?*

It is tempting to try to answer (affirmatively) this question by establishing a suitable spectral gap. Unfortunately the current approaches to the spectral gap use algebraic properties of the matrix coefficients, and thus do not apply over any open set, see [22].

## 2. STATEMENT OF MAIN RESULT

We state our main result, Theorem B below, of which Theorem A is an easy consequence. Initially, we recall the notion of Lyapunov exponents and

Given any compactly supported probability measure  $\nu$  on  $G = \mathrm{GL}(\mathbb{R}^d)$ , let  $\nu^{\mathbb{N}}$  and  $\nu^{\mathbb{Z}}$  denote the corresponding Bernoulli measures on  $G^{\mathbb{N}}$  and  $G^{\mathbb{Z}}$ , respectively. Consider the shift maps  $\sigma : G^{\mathbb{N}} \rightarrow G^{\mathbb{N}}$  and  $\sigma : G^{\mathbb{Z}} \rightarrow G^{\mathbb{Z}}$  given by

$$\sigma((g_n)_n) = (g_{n+1})_n.$$

By the Oseledets multiplicative ergodic theorem (see [65, Theorems 4.1 and 4.2]), there exist  $k \in \{1, \dots, d\}$  and real numbers

$$(2.1) \quad \chi_1(\nu) > \dots > \chi_k(\nu)$$

such that for  $\nu^{\mathbb{N}}$ -almost every  $\mathbf{g} = (g_0, \dots, g_n, \dots) \in G^{\mathbb{N}}$  there exists a decreasing family of vector subspaces

$$(2.2) \quad \mathbb{R}^d = V^1(\mathbf{g}) > \dots > V^k(\mathbf{g}) > V^{k+1}(\mathbf{g}) = \{0\}$$

and for  $\nu^{\mathbb{Z}}$ -almost every  $\mathbf{g}_{\pm} = (\dots, g_{-n}, \dots, g_0, \dots, g_n, \dots) \in G^{\mathbb{Z}}$  there exists a direct sum decomposition

$$(2.3) \quad \mathbb{R}^d = E^1(\mathbf{g}_{\pm}) \oplus \dots \oplus E^k(\mathbf{g}_{\pm})$$

such that, for every  $i = 1, \dots, k$ ,

- $g_0 V^i(\mathbf{g}) = V^i(\sigma(\mathbf{g}))$  and  $g_0 E^i(\mathbf{g}_{\pm}) = E^i(\sigma(\mathbf{g}_{\pm}))$  for  $\nu$ -almost every  $g_0 \in G$ ;
- $V^i(\mathbf{g}) = E^i(\mathbf{g}_{\pm}) \oplus V^{i+1}(\mathbf{g})$  for  $\mathbf{g} = \pi(\mathbf{g}_{\pm})$ , where  $\pi : G^{\mathbb{Z}} \rightarrow G^{\mathbb{N}}$  denotes the canonical projection.
- for every non-zero  $v_i \in V^i(\mathbf{g}) \setminus V^{i+1}(\mathbf{g})$  and  $\nu^{\mathbb{N}}$ -almost every  $\mathbf{g} \in G^{\mathbb{N}}$

$$(2.4) \quad \lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 v_i\| = \chi_i(\nu);$$

- for every non-zero  $v_i \in E^i(\mathbf{g}_{\pm})$  and  $\nu^{\mathbb{Z}}$ -almost every  $\mathbf{g}_{\pm} \in G^{\mathbb{Z}}$

$$(2.5) \quad \lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 v_i\| = \chi_i(\nu) = \lim_n \frac{1}{-n} \log \|g_{-n}^{-1} \cdots g_{-1}^{-1} v_i\|.$$

The maps  $\mathbf{g} \mapsto V^i(\mathbf{g})$  and  $\mathbf{g}_{\pm} \mapsto E^i(\mathbf{g}_{\pm})$  with values in the Grassmannian of  $\mathbb{R}^d$  are measurable and the dimensions  $\dim V^i(\mathbf{g})$  and  $\dim E^i(\mathbf{g}_{\pm})$  are constant on full measure sets. The number  $m_i = \dim V^i - \dim V^{i+1} = \dim E^i$  is called the *multiplicity* of the Lyapunov exponent  $\chi_i(\nu)$ . Denote by  $\lambda_1(\nu) \geq \dots \geq \lambda_d(\nu)$  the Lyapunov exponents *counted with multiplicity*.

Let  $P = \mathbb{P}(\mathbb{R}^d)$ . The *random walk* defined by  $\nu$  is described by the pair  $(F, \nu^{\mathbb{N}})$ , where

$$(2.6) \quad F : G^{\mathbb{N}} \times P \rightarrow G^{\mathbb{N}} \times P, \quad (\mathbf{g}, v) \mapsto (\sigma(\mathbf{g}), g_0 v).$$

The Lyapunov exponents and the Oseledecs filtration may also be obtained from it, as follows. Define

$$(2.7) \quad \Phi : G \times P \rightarrow \mathbb{R}, \quad \Phi(g, v) = \log \frac{\|gv\|}{\|v\|}.$$

(For notational simplicity, we use the same symbol  $(v, \text{ say})$  to denote both a non-zero vector in  $\mathbb{R}^d$  and the corresponding element of  $P$ ; analogously, we use the same notation  $(L, \text{ say})$  for a vector subspace of  $\mathbb{R}^d$  and the subset of  $P$  associated to it.) A result of Ledrappier (see [65, Theorem 6.1]) asserts that:

- Given any  $F$ -invariant ergodic probability measure  $m$  on  $G^{\mathbb{N}} \times P$  that projects to  $\nu^{\mathbb{N}}$ , there exists  $j \in \{1, \dots, k\}$  such that

$$(2.8) \quad \int_{G^{\mathbb{N}} \times P} \Phi \, dm = \chi_j(\nu) \text{ and } m(\{(g, v) : v \in V^j(g) \setminus V^{j+1}(g)\}) = 1.$$

- Given any  $j \in \{1, \dots, k\}$  there is an ergodic  $F$ -invariant probability measure  $m$  projecting to  $\nu^{\mathbb{N}}$  and satisfying (2.8).

Let  $(A_{1,k}, \dots, A_{m,k})$ ,  $k \in \mathbb{N}$  be a sequence of  $m$ -uples of matrices converging to some  $(A_{1,\infty}, \dots, A_{m,\infty}) \in G^m$  and  $(p_{1,k}, \dots, p_{m,k})$ ,  $k \in \mathbb{N}$  be a sequence of probability  $m$ -vectors real numbers converging to some  $(p_{1,\infty}, \dots, p_{m,\infty}) \in \Delta_m$ . Let  $\nu_k$  and  $\nu_\infty$  be the probability measures in  $G$  given by

$$(2.9) \quad \nu_k = \sum_{i=1}^m p_{i,k} \delta_{A_{i,k}} \text{ and } \nu_\infty = \sum_{i=1}^m p_{i,\infty} \delta_{A_{i,\infty}}.$$

We want to prove that  $\lambda_j(\nu_k) \rightarrow \lambda_j(\nu_\infty)$  when  $k \rightarrow \infty$ , for every  $j = 1, \dots, d$ .

In fact, we prove a stronger statement, involving probability measures whose supports need not be finite. Let  $\mathcal{P}_c(G)$  be the space of compactly supported probability measures on  $G$ , with the smallest topology  $\mathcal{T}$  that contains both:

- $\mathcal{W}$  = the restriction of the weak\* topology in the space of probability measures on  $G$
- $\text{supp}^* \mathcal{H}$  = the pull-back under  $\nu \mapsto \text{supp } \nu$  of the Hausdorff topology  $\mathcal{H}$  in the space of compact subsets of  $G$ .

This topology  $\mathcal{T}$  is metrizable, because both  $\mathcal{W}$  and  $\mathcal{H}$  are. A sequence  $(\nu_k)$  converges to  $\nu_\infty$  in  $\mathcal{P}_c(G)$  if and only if

- $(\nu_k)_k \rightarrow \nu_\infty$  in the weak\* topology and
- $(\text{supp } \nu_k)_k \rightarrow \text{supp } \nu_\infty$  in the Hausdorff topology.

That is the case for the measures in (2.9) if  $A_{i,k} \rightarrow A_{i,\infty}$  and  $p_{i,k} \rightarrow p_{i,\infty}$  for every  $i = 1, \dots, m$ ; here, the assumption that  $p_{i,\infty} > 0$  is important to ensure continuity of the supports.

Related to this, the example of Kifer [50] shows that part (i) alone is not enough to ensure continuity of the Lyapunov exponents: for our results to hold one cannot omit part (ii) of the definition of the topology.

**Remark 2.1.** If  $(\nu_k)_k \rightarrow \nu_\infty$  in the weak\* topology then, given any  $\varepsilon > 0$ , the support of  $\nu_\infty$  is contained in the  $\varepsilon$ -neighborhood of  $\text{supp } \nu_k$  for all large  $k$ . Thus, the condition (i) in the definition contains half of the condition (ii). The other half is that, given any  $\varepsilon > 0$ , the support of  $\nu_k$  is contained in the  $\varepsilon$ -neighborhood of  $\text{supp } \nu_\infty$  for all large  $k$ . This will be used repeatedly.

Also, let us point out that for proving Theorem A it suffices to consider the case  $j = 1$ . That is because of the following construction. Let  $E = \mathbb{R}^d$  and  $1 \leq l \leq d$ . The *exterior  $l$ -power*  $\Lambda^l E$  of  $E$  is the vector space of alternating  $l$ -linear forms  $\omega : E^* \times \cdots \times E^* \rightarrow \mathbb{R}$  on the dual space  $E^*$ . The *exterior product* of vectors  $v_1, \dots, v_l \in E$ , is the alternating  $l$ -linear form  $v_1 \wedge \cdots \wedge v_l : E^* \times \cdots \times E^* \rightarrow \mathbb{R}$  defined by

$$(v_1 \wedge \cdots \wedge v_l)(\phi_1, \dots, \phi_l) = \sum_{\sigma} \text{sign}(\sigma) \phi_{\sigma(1)}(v_1) \cdots \phi_{\sigma(l)}(v_l),$$

where the sum is over all permutations of  $\{1, \dots, l\}$ . If  $\{e_j : j = 1, \dots, d\}$  is a basis of  $E$  then  $\{e_{j_1} \wedge \cdots \wedge e_{j_l} : 1 \leq j_1 < \cdots < j_l \leq d\}$  is a basis of  $\Lambda^l E$ . So,

$$\dim \Lambda^l E = \binom{d}{l}.$$

Every  $g \in G$  induces an invertible linear map  $\Lambda^l g : \Lambda^l E \rightarrow \Lambda^l E$ , defined by

$$\Lambda^l g(\omega) : (\phi_1, \dots, \phi_l) \mapsto \omega(\phi_1 \circ g, \dots, \phi_l \circ g),$$

for  $\omega \in \Lambda^l E$  and  $\phi_1, \dots, \phi_l \in E^*$ . Thus, any measure  $\nu$  in  $G$  induces a measure  $\Lambda^l \nu$  in  $\text{GL}(\Lambda^l E)$ , by push-forward under  $g \mapsto \Lambda^l g$ . Moreover, the maps  $\nu \mapsto \Lambda^l \nu$  are continuous. One can check (see [65, Proposition 4.17]) that the Lyapunov exponents of the random walk defined by  $\Lambda^l \nu$ , counted with multiplicity, are the sums

$$\lambda_{i(1)}(\nu) + \cdots + \lambda_{i(l)}(\nu) \text{ with } 1 \leq i_1 < \cdots < i_l \leq d.$$

In particular,

$$\lambda_1(\Lambda^l \nu) = \lambda_1(\nu) + \cdots + \lambda_l(\nu).$$

Thus, proving that  $\nu \mapsto \lambda_1(\Lambda^l \nu)$  is continuous, for every  $1 \leq l \leq d$ , will entail that  $\nu \mapsto \lambda_j(\nu)$  is continuous, for every  $1 \leq j \leq d$ .

In view of these observations, Theorem A will follow immediately from:

**Theorem B.** *The function  $\lambda_1 : \mathcal{P}_c(G) \rightarrow \mathbb{R}$ ,  $\nu \mapsto \lambda_1(\nu)$  is continuous, in any dimension  $d \geq 2$ .*

The rest of the paper is devoted to proving Theorem B. In Sections 3 and 4 we present a useful large deviations principle for Lyapunov exponents (Theorem 4.1). In Sections 5 and 6 we introduce several useful tools. In Section 7 we reduce the proof of Theorem B to a main technical result, Theorem 7.1. The proof of the latter result is by induction on the dimension  $r$  of the equator, as outlined in Section 7. The case  $r = 1$  is carried out in detail in Sections 8 through 11. The inductive step is dealt with in Sections 12 through 16.

Before stating Theorem 7.1 and outlining its proof, we must introduce several general notions and a number of auxiliary results. On the other hand, the proofs of those results, given in Sections 3 to 6, are in themselves not used for establishing Theorem 7.1. Thus the reader is encouraged to skip them at first reading, proceeding as directly as possible to Section 7.

### 3. INVARIANT SUBSPACES

In this section we introduce some background material, due mostly to Furstenberg and Kifer [40, 42]. This also allows us to introduce some notations that will be useful in the following. Proofs and more information can also be found in Chapters 4 through 6 of [65].

**3.1. Stationary measures.** Fix  $\nu \in \mathcal{P}_c(G)$ . We say that a probability measure  $\eta$  on  $P$  is  $\nu$ -stationary if

$$\int_P \psi(x) d\eta(x) = \int_{G \times P} \psi(gy) d\nu(g) d\eta(y)$$

for every bounded measurable function  $\psi : P \rightarrow \mathbb{R}$ . In other words,  $\eta$  is  $\nu$ -stationary if and only if  $\mathcal{P}_\nu^* \eta = \eta$ , where  $\mathcal{P}_\nu^*$  is the operator defined by

$$(3.1) \quad \mathcal{P}_\nu^* \eta = \int_G (g_* \eta) d\nu(g)$$

in the space of probability measures. Moreover (see [65, Proposition 5.5]),  $\eta$  is  $\nu$ -stationary if and only if the probability measure  $\nu^{\mathbb{N}} \times \eta$  is invariant under the projective cocycle  $F : G^{\mathbb{N}} \times P \rightarrow G^{\mathbb{N}} \times P$  defined in (2.6). Stationary measures always exist (see [65, Proposition 5.6]).

We also consider the operator  $\mathcal{P}_\nu : \mathcal{B}(P) \rightarrow \mathcal{B}(P)$  defined in the space  $\mathcal{B}(P)$  of measurable bounded functions  $\psi : P \rightarrow \mathbb{R}$  by

$$(3.2) \quad \mathcal{P}_\nu \psi(v) = \int_G \psi(gv) d\nu(g).$$

A function  $\psi$  is said to be  $\nu$ -stationary if  $\mathcal{P}_\nu \psi = \psi$ . A  $\nu$ -stationary measure  $\eta$  is *ergodic* if every  $\nu$ -stationary function is constant on some full  $\eta$ -measure set. This happens (see [65, Proposition 5.13]), if and only if the  $F$ -invariant measure  $\nu^{\mathbb{N}} \times \eta$  is ergodic for  $F$ . The ergodic decomposition theorem (see [65, Theorem 5.14]) asserts that every  $\nu$ -stationary measure is a convex combination of ergodic  $\nu$ -stationary measures.

Let  $\Phi$  be as in (2.7). Then (see [65, Proposition 6.7]),

$$(3.3) \quad \lambda_1(\nu) = \max \left\{ \int_{G \times P} \Phi d(\nu \times \eta) : \eta \text{ is a } \nu\text{-stationary measure} \right\}.$$

Denote  $\alpha(\eta) = \int_{G \times P} \Phi d(\nu \times \eta)$  for each  $\nu$ -stationary measure  $\eta$ . The ergodic decomposition theorem implies that the maximum does not change if we restrict to ergodic  $\nu$ -stationary measures.

We may also view  $\Phi$  as a function on  $G^{\mathbb{N}} \times P$  that depends only on  $g_0$  and  $v$ :

$$\Phi(\mathbf{g}, v) = \log \frac{\|g_0 v\|}{\|v\|}.$$

Then

$$\lim_n \frac{1}{n} \log \frac{\|g_{n-1} \cdots g_0 v\|}{\|v\|} = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \log \frac{\|g_j g_{j-1} \cdots g_0 v\|}{\|g_{j-1} \cdots g_0 v\|} = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \Phi(F^j(\mathbf{g}, v)).$$

So, for any ergodic  $\nu$ -stationary measure  $\eta$ , we have

$$(3.4) \quad \lim_n \frac{1}{n} \log \frac{\|g_{n-1} \cdots g_0 v\|}{\|v\|} = \int_{G \times P} \Phi d(\nu \times \eta)$$

$\nu^{\mathbb{N}} \times \eta$ -almost everywhere in  $G^{\mathbb{N}} \times P$ .

Furstenberg and Kifer have shown (see [42, Theorem 2.1]) that if  $\alpha(\eta) = \lambda_1(\nu)$  for every (ergodic)  $\nu$ -stationary measure  $\eta$  then for every  $v \in P$

$$(3.5) \quad \lim_n \frac{1}{n} \log \frac{\|g_{n-1} \cdots g_0 v\|}{\|v\|} = \lambda_1(\nu) \text{ for } \nu^{\mathbb{N}}\text{-almost every } \mathbf{g}.$$

This fact is also contained in Theorem 4.1 below. In the next section we analyze what happens when the hypothesis of (3.5) is not fulfilled.

**3.2. The equator.** A vector subspace  $L$  of  $\mathbb{R}^d$  is said to be  $\nu$ -invariant if  $gL = L$  for  $\nu$ -almost every  $g$  or, equivalently, for every  $g \in \text{supp } \nu$ . Observe that if  $L$  is a  $\nu$ -invariant subspace then

$$\lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 | L\|$$

exists for every  $\mathbf{g}$  in a full  $\nu^{\mathbb{N}}$ -measure subset of  $G^{\mathbb{N}}$  and is constant on this subset. This direct consequence of Kingman's subadditive ergodic theorem (see [65, Theorem 3.3]), together with the fact that the Bernoulli shift  $(\sigma, \nu^{\mathbb{N}})$  is ergodic, will be used repeatedly. It is part of the proof of the Oseledets theorem (see [65, Proposition 4.11]) that for  $L = \mathbb{R}^d$  the limit coincides with the largest Lyapunov  $\lambda_1(\nu)$ .

For any ergodic  $\nu$ -stationary measure  $\eta$ , define  $L(\eta)$  to be the smallest vector subspace such that  $\eta(L(\eta)) = 1$ . Equivalently,  $L(\eta)$  is the vector subspace spanned by the support of  $\eta$ . Then  $L(\eta)$  is  $\nu$ -invariant:

$$\int_{G \times P} \chi_{L(\eta)}(gv) d\nu(g) d\eta(v) = \int_P \chi_{L(\eta)}(v) d\eta(v) = \eta(L) = 1$$

and this implies

$$\eta(g^{-1}L(\eta)) = \int_P \chi_{L(\eta)}(gv) d\eta(v) = 1 \text{ for } \nu\text{-almost every } g.$$

So, it follows from the definition that  $g^{-1}L(\eta) = L(\eta)$  for  $\nu$ -almost every  $g$ .

Since the support of  $\eta$  spans  $L(\eta)$ , it follows from (3.4) that we may find a basis  $v_1, \dots, v_l$  of  $L(\eta)$  and a full  $\nu^{\mathbb{N}}$ -measure subset of  $\mathbf{g}$  such that

$$\lim_n \frac{1}{n} \log \frac{\|g_{n-1} \cdots g_0 v_i\|}{\|v_i\|} = \alpha(\eta) \text{ for } i = 1, \dots, l.$$

Then

$$(3.6) \quad \lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 | L(\eta)\| = \alpha(\eta).$$

It also follows that if  $\eta$  is such that  $\alpha(\eta) < \lambda_1(\nu)$  then  $L(\eta)$  is a proper subspace.

We call the *equator* of  $\nu$  a maximal  $\nu$ -invariant subspace  $E$  such that

$$(3.7) \quad \lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 | E\| < \lambda_1(\nu).$$

Such a subspace is necessarily proper, but it may not exist. It follows from the previous paragraph that the equator does exist if  $\alpha(\eta) < \lambda_1(\nu)$  for some (ergodic)  $\nu$ -stationary measure; the converse is also true. Moreover, the equator is unique and contains  $L(\eta)$  for every ergodic  $\nu$ -stationary measure  $\eta$ : both claims follow immediately from the observation that if two subspaces satisfy (3.7) then so does their sum.

If the equator does exist, every  $g$  in (the group generated by) the support of  $\nu$  may be written as

$$(3.8) \quad g = \begin{pmatrix} g^E & h \\ 0 & g^\perp \end{pmatrix} \text{ with } g^E \in \text{GL}(E) \text{ and } g^\perp \in \text{GL}(E^\perp).$$

Let  $\nu^E$  and  $\nu^\perp$  be the push-forwards of  $\nu$  under the maps  $g \mapsto g^E$  and  $g \mapsto g^\perp$ . Furstenberg and Kifer [42, Lemma 3.6] observed that

$$\lambda_1(\nu) = \max\{\lambda_1(\nu^E), \lambda_1(\nu^\perp)\}.$$

The property (3.7) means that  $\lambda_1(\nu^E) < \lambda_1(\nu)$ . Hence,  $\lambda_1(\nu^\perp) = \lambda_1(\nu)$ .

**Remark 3.1.** Any  $g \in G$  that preserves  $E$  may be written in the form (3.8). Then  $(gu)^\perp = g^\perp u^\perp$  for every  $u$ , where  $\perp$  denotes the component orthogonal to  $E$ . Do not mistake this for  $gu^\perp = hu^\perp + g^\perp u^\perp$ . It is equally clear that  $\|g^\perp\| \leq \|g\|$ . These simple facts will be used several times.

Suppose that there exists some ergodic  $\nu^\perp$ -stationary measure  $\eta^\perp$  such that

$$\alpha^\perp(\eta^\perp) = \int_{G \times P} \Phi d(\nu^\perp \times \eta^\perp)$$

is strictly less than  $\lambda_1(\nu^\perp) = \lambda_1(\nu)$ . Then  $L(\eta^\perp)$  is a proper  $\nu^\perp$ -invariant subspace of  $E^\perp$  and, according to (3.6),

$$\lim_n \frac{1}{n} \log \|g_{n-1}^\perp \cdots g_0^\perp | L(\eta^\perp)\| = \alpha^\perp(\eta^\perp).$$

Then  $E' = E \oplus L(\eta^\perp)$  is a proper  $\nu$ -invariant subspace of  $\mathbb{R}^d$  and it satisfies the equator property (3.7), because

$$\begin{aligned} & \lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 | E \oplus L(\eta^\perp)\| \\ & \leq \max \left\{ \lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 | E\|, \lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 | L(\eta^\perp)\| \right\} \end{aligned}$$

Since  $E'$  contains  $E$  strictly, this contradicts the definition of the equator. This contradiction proves that  $\alpha^\perp(\eta^\perp) = \lambda_1(\nu)$  for every  $\nu^\perp$ -stationary measure  $\eta^\perp$ .

So, by (3.5), for every  $v^\perp \in \mathbb{P}(E^\perp)$  there exists a full  $\nu^\mathbb{N}$ -measure subset of  $\mathbf{g}$  for which

$$(3.9) \quad \lim_n \frac{1}{n} \log \frac{\|g_{n-1}^\perp \cdots g_0^\perp v^\perp\|}{\|v^\perp\|} = \lambda_1(\nu).$$

This also implies (see [65, Proposition 4.14]) that for every  $v \in \mathbb{R}^d \setminus E$  there exists a full  $\nu^\mathbb{N}$ -measure subset of  $\mathbf{g}$  for which

$$(3.10) \quad \lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 v\| = \lim_n \frac{1}{n} \log \frac{\|g_{n-1} \cdots g_0 v\|}{\|v\|} = \lambda_1(\nu).$$

The following example shows that the equator need not be any of the subspaces  $V^i$  in the Oseledets flag (2.2):

**Example 3.2.** Let  $d = 3$  and  $m = 2$  and  $\nu = p_1 \delta_{A_1} + p_2 \delta_{A_2}$  where the matrices  $A_1$  and  $A_2$  are given by

$$\begin{aligned} A_1 &= \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix} \text{ with } B = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma^{-1} \end{pmatrix} \text{ and } \sigma > 1 \\ A_2 &= \begin{pmatrix} R_\theta B & 0 \\ 0 & 1 \end{pmatrix} \text{ with } R_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \text{ and } \theta \neq 0 \text{ small.} \end{aligned}$$

The subspace  $E_3 = \{(0, 0)\} \times \mathbb{R}$  is  $\nu$ -invariant and corresponds to a zero Lyapunov exponent. The only other  $\nu$ -invariant subspace is  $E_{12} = \mathbb{R}^2 \times \{0\}$ . Given any  $\varepsilon > 0$ , the cone

$$C^u = \{(x, y, 0) \in \mathbb{R}^3 : |y| \leq \varepsilon|x\} \subset E_{12}$$

is forward invariant under both  $A_1$  and  $A_2$ , as long as  $\theta$  is close enough to zero. This implies that some Lyapunov exponent is close to  $\log \sigma$ . Analogously, the cone

$$C^s = \{(x, y, 0) \in \mathbb{R}^3 : |x| \leq \varepsilon|y|\} \subset E_{12}$$

is backward invariant under both  $A_1$  and  $A_2$ , which implies that some Lyapunov exponent is close to  $-\log \sigma$ . So, the Lyapunov exponents are

$$\lambda_1(\nu) \approx \log \sigma \text{ and } \lambda_2(\nu) = 0 \text{ and } \lambda_3(\nu) \approx -\log \sigma.$$

Thus,  $E_3$  is the equator of  $\nu$  and corresponds to the middle eigenvalue  $\lambda_2(\nu) = 0$ . Note that both  $A_1^\perp = B$  and  $A_2^\perp = R_\theta B$  have determinant 1.

#### 4. UNIFORM CONVERGENCE IN MEASURE

We need to prove that the limit in (3.9) is *uniform in measure* with respect to  $v \in \mathbb{P}(E^\perp)$ . This follows directly from a corresponding fact for the limit in (3.5), that we state precisely as follows:

**Theorem 4.1.** *Assume that  $\alpha(\eta) = \lambda_1(\nu)$  for every  $\nu$ -stationary measure  $\eta$ . Then for any  $\varepsilon > 0$  there exist constants  $C = C(\nu, \varepsilon) > 0$  and  $c = c(\nu, \varepsilon) > 0$  such that for any  $v \in P$  and  $N \in \mathbb{N}$ , there exists a measurable set  $\mathcal{E} = \mathcal{E}(\nu, \varepsilon, v, N) \subset G^\mathbb{N}$  satisfying:*

- (1)  $\nu^\mathbb{N}(\mathcal{E}^c) \leq Ce^{-cN}$  and
- (2)  $\frac{1}{n} \log \frac{\|g_{n-1} \cdots g_0 v\|}{\|v\|} \in (\lambda_1(\nu) - \varepsilon, \lambda_1(\nu) + \varepsilon)$  for every  $\mathbf{g} \in \mathcal{E}$  and  $n \geq N$ .

In what follows we prove Theorem 4.1. The proof will not be used in the rest of the paper, so the reader may choose to skip the remainder of this section at first reading.

Recall that  $S = \text{supp } \nu$  is taken to be compact. Let  $C(S \times P)$  be the Banach space of continuous functions  $\psi : S \times P \rightarrow \mathbb{R}$  with the norm  $\|\psi\| = \sup |\psi|$ . Consider the operator  $\mathcal{Q}_\nu$  defined on  $C(S \times P)$  by

$$\mathcal{Q}_\nu \psi(g', v) = \int_S \psi(g, gv) d\nu(g).$$

Note that  $\mathcal{Q}_\nu \psi(g', v)$  does not depend on  $g'$ . The dual operator  $\mathcal{Q}_\nu^*$  acts in the space of finite signed measures on  $S \times P$  by

$$\int_{S \times P} \psi d\mathcal{Q}_\nu^* \lambda = \int_{S \times P} \mathcal{Q}_\nu \psi d\lambda \text{ for every } \psi \in C(S \times P).$$

A signed measure  $\lambda$  is said to be  $\mathcal{Q}_\nu$ -invariant if  $\mathcal{Q}_\nu^* \lambda = \lambda$ . Given  $\varphi \in C(S \times P)$ , denote

$$\Sigma(\nu, \varphi) = \sup \left\{ \int_{S \times P} \varphi d\lambda : \lambda \text{ is a } \mathcal{Q}_\nu\text{-invariant probability measure} \right\}.$$

**Lemma 4.2.** *Given any non-negative  $\varphi \in C(S \times P)$  and any  $\varepsilon > 0$  there exist  $\psi, \xi \in C(S \times P)$  such that  $\varphi = \mathcal{Q}_\nu \psi - \psi + \xi$  and  $\|\xi\| \leq \Sigma(\nu, \varphi) + \varepsilon$ .*

*Proof.* Let  $\mathcal{W} \subset C(S \times P)$  be the subspace of functions of the form  $\varphi = \mathcal{Q}_\nu \psi - \psi$  for some  $\psi \in C(S \times P)$ . By Hahn–Banach, given any  $\varphi \in C(S \times P)$  there exists a continuous linear functional  $L : C(S \times P) \rightarrow \mathbb{R}$  such that  $\|L\| = 1$  and  $L|_{\mathcal{W}} \equiv 0$  and  $L(\varphi) = d(\varphi, \mathcal{W})$ . By Riesz–Markov, there exists some signed measure  $\lambda$  on

$S \times P$  such that  $\|\lambda\| = 1$  and  $L(\psi) = \int_{S \times P} \psi d\lambda$  for every  $\psi \in C(S \times P)$ . The fact that  $L$  vanishes on  $W$  ensures that  $\mathcal{Q}_\nu^* \lambda = \lambda$ :

$$\int_{S \times P} (\mathcal{Q}_\nu \psi - \psi) d\lambda = L(\mathcal{Q}_\nu \psi - \psi) = 0 \text{ for every } \psi \in C(S \times P).$$

Let  $\lambda = \lambda^+ - \lambda^-$  be the Hahn decomposition of  $\lambda$ . Then  $\|\lambda^+\| + \|\lambda^-\| = \|\lambda\| = 1$  and  $\lambda = \mathcal{Q}_\nu^* \lambda = \mathcal{Q}_\nu^* \lambda^+ - \mathcal{Q}_\nu^* \lambda^-$ . Since  $\mathcal{Q}_\nu^* \lambda^\pm$  are non-negative measures, the latter implies that  $\mathcal{Q}_\nu^* \lambda^\pm \geq \lambda^\pm$ . Since  $\mathcal{Q}_\nu 1 = 1$  we have

$$\|\mathcal{Q}_\nu^* \lambda^\pm\| = \int_{S \times P} 1 d\mathcal{Q}_\nu^* \lambda^\pm = \int_{S \times P} \mathcal{Q}_\nu 1 d\lambda^\pm = \int_{S \times P} 1 d\lambda^\pm = \|\lambda^\pm\|.$$

Hence,  $\mathcal{Q}_\nu^* \lambda^\pm = \lambda^\pm$ . Let  $\lambda_0 = \lambda^+ / \|\lambda^+\|$ . Then  $\mathcal{Q}_\nu^* \lambda_0 = \lambda_0$  and  $\lambda_0 \geq \lambda^+ \geq \lambda$ . Since  $\varphi$  is assumed to be non-negative, it follows that

$$(4.1) \quad \int_{S \times P} \varphi d\lambda_0 \geq \int_{S \times P} \varphi d\lambda = L(\varphi) = d(\varphi, \mathcal{W})$$

Take  $\psi \in C(S \times P)$  such that  $\|\varphi - (\mathcal{Q}_\nu \psi - \psi)\| \leq d(\varphi, \mathcal{W}) + \varepsilon$  and then define  $\xi = \varphi - (\mathcal{Q}_\nu \psi - \psi)$ . Then, by (4.1),

$$\|\xi\| \leq d(\varphi, \mathcal{W}) + \varepsilon \leq \int_{S \times P} \varphi d\lambda_0 + \varepsilon$$

and this implies that  $\|\xi\| \leq \Sigma(\nu, \varphi) + \varepsilon$ .  $\square$

For  $n \geq 0$ , let  $\mathcal{B}_n$  be the  $\sigma$ -algebra of  $S^\mathbb{N}$  generated by the family of cylinders  $[0; A_0, \dots, A_n]$ , where the  $A_j$  are measurable subsets of  $S$ . For a measurable function  $X : S^\mathbb{N} \rightarrow \mathbb{R}$ , let  $\mathbb{E}(X) = \int_{G^\mathbb{N}} X d\nu^\mathbb{N}$  and  $\mathbb{E}(X | \mathcal{B}_n)$  denote the *expectation* of  $X$  conditioned to  $\mathcal{B}_n$ , that is, the (essentially unique)  $\mathcal{B}_n$ -measurable function such that

$$\int_{G^\mathbb{N}} \mathbb{E}(X | \mathcal{B}_n) d\nu^\mathbb{N} = \int_{G^\mathbb{N}} X d\nu^\mathbb{N}.$$

The next lemma is a particular instance of the Azuma–Hoeffding inequality (Azuma [11], Hoeffding [48]) for sums of bounded random variables:

**Lemma 4.3.** *Let  $Y_n : G^\mathbb{N} \rightarrow \mathbb{R}$ ,  $n \in \mathbb{N}$  be such that  $A = \sup_n \|Y_n\|$  is finite and*

$$(4.2) \quad \mathbb{E}(Y_{i_1} \cdots Y_{i_k}) = 0 \text{ for every } 1 \leq i_1 < \cdots < i_k.$$

*Then, for any  $s > 0$  and  $n \in \mathbb{N}$ ,*

$$\nu^\mathbb{N}(\{\mathfrak{g} : |Y_1 + \cdots + Y_n|(\mathfrak{g}) \geq s\}) \leq 2 \exp\left(-\frac{s^2}{2nA^2}\right).$$

*Proof.* Since the exponential function is convex,

$$e^{ax} = \exp\left(a\left(\frac{1+x}{2}\right) - a\left(\frac{1-x}{2}\right)\right) \leq \frac{1+x}{2}e^a + \frac{1-x}{2}e^{-a} = \cosh a + x \sinh a$$

for every  $x \in [-1, 1]$  and  $a \in \mathbb{R}$ . Taking  $x = Y_i / \|Y_i\|$  and  $a = t\|Y_i\|$ , we get that

$$e^{tY_i} \leq \cosh(t\|Y_i\|) + \frac{Y_i}{\|Y_i\|} \sinh(t\|Y_i\|)$$

for any  $t \in \mathbb{R}$ . The hypothesis (4.2) implies that

$$\mathbb{E}\left(\prod_{i=1}^n (a_i + b_i Y_i)\right) = \prod_{i=1}^n a_i$$



for any real numbers  $a_1, \dots, a_n$  and  $b_1, \dots, b_n$ . Hence,

$$\mathbb{E} \left( e^{t \sum_{i=1}^n Y_i} \right) \leq \mathbb{E} \left( \prod_{i=1}^n \cosh(t \|Y_i\|) + \frac{Y_i}{\|Y_i\|} \sinh(t \|Y_i\|) \right) = \prod_{i=1}^n \cosh(t \|Y_i\|).$$

Using the fact that, for every  $x \in \mathbb{R}$ ,

$$\cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \leq \sum_{k=0}^{\infty} \frac{x^{2k}}{2^k k!} = e^{x^2/2}$$

we conclude that

$$\mathbb{E} \left( e^{t \sum_{i=1}^n Y_i} \right) \leq \prod_{i=1}^n \exp \left( \frac{t^2}{2} \|Y_i\|^2 \right) = \exp \left( \frac{t^2}{2} \sum_{i=1}^n \|Y_i\|^2 \right)$$

for any  $t \in \mathbb{R}$ . Then, by the Chebyshev inequality,

$$\begin{aligned} \nu^{\mathbb{N}} \left( \left\{ \mathbf{g} : \sum_{j=1}^n Y_j(\mathbf{g}) \geq s \right\} \right) &= \nu^{\mathbb{N}} \left( \left\{ \mathbf{g} : e^{t \sum_{j=1}^n Y_j(\mathbf{g})} \geq e^{ts} \right\} \right) \\ &\leq e^{-ts} \mathbb{E} \left( e^{t \sum_{j=1}^n Y_j} \right) \leq \exp \left( -ts + \frac{t^2}{2} \sum_{i=1}^n \|Y_i\|^2 \right) \end{aligned}$$

for any  $t \in \mathbb{R}$ . Taking  $t = s / \sum_{i=1}^n \|Y_i\|^2$ , we conclude that

$$\nu^{\mathbb{N}} \left( \left\{ \mathbf{g} : \sum_{i=1}^n Y_j(\mathbf{g}) \geq s \right\} \right) \leq \exp \left( -\frac{s^2}{2 \sum_{i=1}^n \|Y_i\|^2} \right) \leq \exp \left( -\frac{s^2}{2nA^2} \right)$$

Analogously, replacing each  $Y_j$  with  $-Y_j$ ,

$$\nu^{\mathbb{N}} \left( \left\{ \mathbf{g} : \sum_{i=1}^n Y_j(\mathbf{g}) \leq -s \right\} \right) \leq \exp \left( -\frac{s^2}{2nA^2} \right).$$

Adding these two inequalities, we get the conclusion of the lemma.  $\square$

**Lemma 4.4.** *For any  $\psi \in C(S \times P)$  and  $\varepsilon > 0$  there exist  $C_1 = C_1(\psi, \varepsilon) > 0$  and  $c_1 = c_1(\psi, \varepsilon) > 0$  such that for any  $v \in \mathbb{R}^d$  and  $N \in \mathbb{N}$  there exists a measurable set  $\mathcal{E}_1 = \mathcal{E}_1(\nu, \psi, \varepsilon, v, N)$  satisfying:*

- (1)  $\nu^{\mathbb{N}}(\mathcal{E}_1^c) \leq C_1 e^{-c_1 N}$  and
- (2)  $\frac{1}{n} \sum_{j=0}^{n-1} \mathcal{Q}_\nu \psi(g_{j-1}, g_{j-1} \cdots g_0 v) - \psi(g_j, g_j \cdots g_0 v) \in (-\varepsilon, \varepsilon)$  for any  $\mathbf{g} \in \mathcal{E}_1$  and  $n \geq N$ .

*Proof.* We may suppose that  $\|\psi\| > 0$  for otherwise the statement is trivial. Given  $v \in \mathbb{R}^d$  and  $n \in \mathbb{N}$ , define

$$Y_n(\mathbf{g}) = \mathcal{Q}_\nu \psi(g_{n-1}, g_{n-1} \cdots g_0 v) - \psi(g_n, g_n \cdots g_0 v).$$

Clearly,  $\|Y_n\| \leq 2\|\psi\|$ . Moreover,  $Y_n$  is  $\mathcal{B}_n$ -measurable and, by the definition of  $\mathcal{Q}_\nu$ ,

$$\mathbb{E}(Y_n \mid \mathcal{B}_{n-1}) = \mathcal{Q}_\nu \psi(g_{n-1}, g_{n-1} \cdots g_0 v) - \int_G \psi(g, g g_{n-1} \cdots g_0 v) d\nu(g) \equiv 0.$$

Thus, given any  $k$ -uple  $(i_1, \dots, i_{k-1}, i_k)$  with  $i_k > i_j$  for every  $j = 1, \dots, k-1$ ,

$$\mathbb{E}(Y_{i_1} \cdots Y_{i_{k-1}} Y_{i_k}) = \mathbb{E}(Y_{i_1} \cdots Y_{i_{k-1}} \mathbb{E}(Y_{i_k} \mid \mathcal{B}_{i_{k-1}})) = 0.$$

So,  $(Y_n)_n$  satisfies the hypotheses of Lemma 4.4. Thus,

$$\nu^{\mathbb{N}} \left( \left\{ \mathbf{g} \in G^{\mathbb{N}} : \frac{1}{n} |Y_0(\mathbf{g}) + \cdots + Y_{n-1}(\mathbf{g})| \geq \varepsilon \right\} \right) \leq 2 \exp \left( -\frac{n\varepsilon^2}{8\|\psi\|^2} \right)$$

for every  $\varepsilon > 0$  and every  $n \in \mathbb{N}$ . The conclusion follows by letting  $\mathcal{E}_1$  be the set of all  $\mathbf{g}$  such that the claim in part (2) holds for every  $n \geq N$  and taking  $c_1 = \varepsilon^2/(8\|\psi\|^2)$  and  $C_1 = 2 \sum_{j=0}^{\infty} e^{-c_1 j}$ .  $\square$

**Lemma 4.5.** *For every  $\varphi \in C(S \times P)$  and  $\varepsilon > 0$  there exist  $C_2 = C_2(\nu, \varphi, \varepsilon) > 0$  and  $c_2 = c_2(\nu, \varphi, \varepsilon) > 0$  such that, for every  $v \in P$  and  $N \in \mathbb{N}$  there exists a measurable set  $\mathcal{E}_2 = \mathcal{E}_2(\nu, \varphi, \varepsilon, v, N) \subset G^{\mathbb{N}}$  satisfying:*

- (1)  $\nu^{\mathbb{N}}(\mathcal{E}_2^c) \leq C_2 e^{-c_2 N}$  and
- (2)  $\frac{1}{n} \sum_{j=0}^{n-1} \varphi(g_j, g_j \dots g_0 v) < \Sigma(\nu, \varphi) + \varepsilon$  for every  $\mathbf{g} \in \mathcal{E}_2$  and  $n \geq N$ .

*Proof.* Neither the hypothesis nor the conclusion are affected if one replaces  $\varphi$  with  $\varphi + \text{const}$  (clearly,  $\Sigma_{\varphi + \text{const}} = \Sigma(\nu, \varphi) + \text{const}$ ). Thus, it is no restriction to suppose that  $\varphi$  is non-negative. Then, by Lemma 4.2, for any  $\varepsilon > 0$  we may find  $\psi, \xi \in C(S \times P)$  such that  $\varphi = \mathcal{Q}_\nu \psi - \psi + \xi$  and  $\|\xi\| < \Sigma(\nu, \varphi) + \varepsilon/4$ . Then

$$\begin{aligned} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(g_j, g_j \dots g_0 v) &= \frac{1}{n} \sum_{j=1}^n \mathcal{Q}_\nu \psi(g_{j-1}, g_{j-1} \dots g_0 v) - \psi(g_j, g_j \dots g_0 v) \\ &\quad + \frac{1}{n} \sum_{j=0}^{n-1} \xi(g_j, g_j \dots g_0 v) + \frac{1}{n} \psi(g_n, g_n \dots g_0 v) - \frac{1}{n} \psi(g_0, g_0 v). \end{aligned}$$

Let  $C_1, c_1$  and  $\mathcal{E}_1$  be as in Lemma 4.3, with  $\varepsilon$  replaced with  $\varepsilon/4$ . The first term in the previous sum is less than  $\varepsilon/4$  for every  $\mathbf{g} \in \mathcal{E}_1$ . The second term is bounded above by  $\|\xi\| < \Sigma(\nu, \varphi) + \varepsilon/4$ . Let  $L = \lceil \|\psi\|/(4\varepsilon) \rceil$ . The third and fourth terms are less than  $\varepsilon/4$  if  $n > L$ . Thus, for every  $\mathbf{g} \in \mathcal{E}_1$  and  $n \geq N > L$ ,

$$\frac{1}{n} \sum_{j=0}^{n-1} \varphi(g_j, g_j \dots g_0 v) < \Sigma(\nu, \varphi) + \varepsilon.$$

Take  $c_2 = c_1$  and  $C_2 = \max\{C_1, e^{c_2 L}\}$ . Define  $\mathcal{E}_2 = \mathcal{E}_1$  if  $N > L$  and  $\mathcal{E}_2 = G^{\mathbb{N}}$  otherwise.  $\square$

**Corollary 4.6.** *Let  $\varphi \in C(S \times P)$  be such that  $\int_{S \times P} \varphi d\lambda = \Sigma(\nu, \varphi)$  for every  $\mathcal{Q}_\nu$ -invariant probability measure  $\lambda$ . For  $\varepsilon > 0$  there exist  $C_3 = C_3(\nu, \varphi, \varepsilon) > 0$  and  $c_3 = c_3(\nu, \varphi, \varepsilon) > 0$  such that, for every  $v \in P$  and  $N \in \mathbb{N}$  there exists a measurable set  $\mathcal{E}_3 = \mathcal{E}_3(\nu, \varphi, \varepsilon, v, N) \subset G^{\mathbb{N}}$  satisfying:*

- (1)  $\nu^{\mathbb{N}}(\mathcal{E}_3^c) \leq C_3 e^{-c_3 N}$  and
- (2)  $\frac{1}{n} \sum_{j=0}^{n-1} \varphi(g_j, g_j \dots g_0 v) \in (\Sigma(\nu, \varphi) - \varepsilon, \Sigma(\nu, \varphi) + \varepsilon)$  for  $\mathbf{g} \in \mathcal{E}_3$  and  $n \geq N$ .

*Proof.* The assumption implies that  $\Sigma(\nu, -\varphi) = -\Sigma(\nu, \varphi)$ . So, applying Lemma 4.5 also to the function  $-\varphi$  we get  $C_2(-\varphi, \varepsilon)$  and  $c_2(-\varphi, \varepsilon)$  and  $\mathcal{E}_2(-\varphi, \varepsilon, v, N)$  such that

$$\frac{1}{n} \sum_{j=0}^{n-1} \varphi(g_j, g_j \dots g_0 v) > \Sigma(\nu, \varphi) - \varepsilon.$$

for any  $\mathbf{g} \in \mathcal{E}_2(-\varphi, \varepsilon, v, N)$  and  $n \geq N$ . To conclude, take  $C_3$  to be the sum of  $C_2(\nu, \pm\varphi, \varepsilon)$  and  $c_3$  to be the minimum of  $c_2(\nu, \pm\varphi, \varepsilon)$  and  $\mathcal{E}_3$  to be the intersection of  $\mathcal{E}_2(\nu, \pm\varphi, \varepsilon, v, N)$ .  $\square$

We also need to interpret  $\Sigma(\nu, \varphi)$  in terms of the  $\nu$ -stationary measures. That is the purpose of the next lemma:

**Lemma 4.7.** *Let  $\iota : S \times P \rightarrow S \times P$ ,  $\iota(g, v) = (g, gv)$ . Then  $\eta \mapsto \iota_* (\nu \times \eta)$  maps the set of  $\nu$ -stationary measures bijectively to the set of  $\mathcal{Q}_\nu$ -invariant probability measures. Its inverse is the push-forward  $\pi_*$  of the canonical projection  $\pi : S \times P \rightarrow P$ .*

*Proof.* Given any probability measure  $\lambda$  with  $\mathcal{Q}_\nu^* \lambda = \lambda$ , let  $\eta$  be its projection to  $P$ . For any bounded measurable function  $\varphi : S \times P \rightarrow \mathbb{R}$ ,

$$\begin{aligned} \int_{S \times P} \varphi(g', v) d\lambda(g', v) &= \int_{S \times P} \mathcal{Q}_\nu \varphi(g', v) d\lambda(g', v) \\ &= \int_{S \times S \times P} \varphi(g, gv) d\nu(g) d\lambda(g', v) = \int_{S \times P} \varphi(g, gv) d\nu(g) d\eta(v) \end{aligned}$$

(because the integrand does not depend on  $g'$ ). Moreover, the special case when  $\varphi$  does not depend on  $g$  means that

$$\begin{aligned} \int_P \phi(v) d\eta(v) &= \int_{S \times P} \phi(v) d\lambda(g', v) = \int_{S \times P} \mathcal{Q}_\nu \phi(g', v) d\lambda(g', v) \\ &= \int_{S \times S \times P} \phi(gv) d\nu(g) d\lambda(g', v) = \int_{S \times P} \phi(gv) d\nu(g) d\eta(v) \end{aligned}$$

for any bounded measurable function  $\phi : P \rightarrow \mathbb{R}$ . In other words,  $\eta$  is  $\nu$ -stationary.

Conversely, given any  $\nu$ -stationary measure  $\eta$ , define  $\lambda$  on  $S \times P$  by

$$\int_{S \times P} \varphi(g', v) d\lambda(g', v) = \int_{S \times P} \varphi(g, gv) d\nu(g) d\eta(v).$$

for any bounded measurable function  $\varphi : S \times P \rightarrow \mathbb{R}$ . Then

$$\begin{aligned} \int_{S \times P} \mathcal{Q}_\nu \varphi(g', v) d\lambda(g', v) &= \int_{S \times P} \mathcal{Q}_\nu \varphi(g', g'v) d\nu(g') d\eta(v) \\ &= \int_{S \times S \times P} \varphi(g, gg'v) d\nu(g) d\nu(g') d\eta(v). \end{aligned}$$

Since  $\eta$  is  $\nu$ -stationary, the right-hand side may be rewritten as

$$\int_{S \times P} \varphi(g, gv) d\nu(g) d\eta(v) = \int_{S \times P} \varphi(g, v) d\lambda(g, v).$$

Combining these two identities, one sees that  $\mathcal{Q}_\nu^* \lambda = \lambda$ .  $\square$

*Proof of Theorem 4.1.* Consider  $\varphi = \Phi \circ \iota^{-1}$ , that is,

$$\varphi : S \times P \rightarrow \mathbb{R}, \quad \varphi(g, v) = \log \frac{\|v\|}{\|g^{-1}v\|}.$$

Lemma 4.7 implies that

$$\begin{aligned} \Sigma(\nu, \varphi) &= \sup \left\{ \int_{S \times P} \varphi d\lambda : \mathcal{Q}_\nu^* \lambda = \lambda \right\} \\ &= \sup \left\{ \int_{S \times P} \Phi d(\nu \times \eta) : \eta \text{ is a } \nu\text{-stationary measure} \right\} = \lambda_1(\nu) \end{aligned}$$

and the hypothesis of Corollary 4.6 corresponds precisely to the hypothesis of Theorem 4.1. Take  $C = C_3(\nu, \varphi, \varepsilon)$  and  $c = c_3(\nu, \varphi, \varepsilon)$  and  $\mathcal{E} = \mathcal{E}_3(\nu, \varphi, \varepsilon, v, N)$ . Then  $\nu^{\mathbb{N}}(\mathcal{E}^c) \leq C e^{-cn}$  and for any  $\mathbf{g} \in \mathcal{E}$ ,

$$\begin{aligned} \frac{1}{n} \log \frac{\|g_{n-1} \cdots g_0 v\|}{\|v\|} &= \frac{1}{n} \sum_{j=0}^{n-1} \Phi(g_j, g_{j-1} \cdots g_0 v) = \frac{1}{n} \sum_{j=0}^{n-1} \varphi(g_j, g_j g_{j-1} \cdots g_0 v) \\ &\in (\Sigma(\nu, \varphi) - \varepsilon, \Sigma(\nu, \varphi) + \varepsilon) = (\lambda_1(\nu) - \varepsilon, \lambda_1(\nu) + \varepsilon). \end{aligned}$$

This completes the proof of the theorem.  $\square$

## Part II. Preliminaries and outline of the proof

### 5. THE EQUATOR IS A REPELLER

Let  $(\nu_k)_k$  be a sequence converging to some  $\nu_\infty$  in the space  $\mathcal{P}_c(G)$ . By (3.3), for every  $k$  there exists a  $\nu_k$ -stationary measure  $\eta_k$  on  $P$  such that

$$\lambda_1(\nu_k) = \int_{G \times P} \Phi d(\nu_k \times \eta_k).$$

Since the space of probability measures on  $P$  is weak\*-compact, to prove that  $(\lambda_1(\nu_k))$  converges to  $\lambda_1(\nu_\infty)$  it is no restriction to suppose that the sequence  $(\eta_k)_k$  converges to some probability measure  $\eta_\infty$ . Then

$$\lambda_1(\nu_k) \rightarrow \int_{G \times P} \Phi d(\nu_\infty \times \eta_\infty) \text{ when } k \rightarrow \infty.$$

The measure  $\eta_\infty$  is necessarily  $\nu_\infty$ -stationary (see [65, Proposition 5.9]).

Now there are two alternatives. If  $\int_{G \times P} \Phi d(\nu_\infty \times \eta_\infty) = \lambda_1(\nu_\infty)$  then

$$\lambda_1(\nu_k) \rightarrow \lambda_1(\nu_\infty),$$

as we wanted to prove. Otherwise, we are in the setting of Section 4: there is a proper subspace  $E$  of  $\mathbb{R}^d$  such that

(i)  $E$  is  $\nu_\infty$ -invariant and

$$\lim \frac{1}{n} \log \|g_{n-1} \cdots g_0 | E\| < \lambda_1(\nu_\infty).$$

(ii) For any  $\varepsilon > 0$  and  $\delta > 0$  there exists  $N = N(\nu_\infty, \varepsilon, \delta) \in \mathbb{N}$  and for every  $v^\perp \in \mathbb{P}(E^\perp)$  there exists  $\mathcal{E} = \mathcal{E}(\nu_\infty, \varepsilon, \delta, v^\perp)$  such that  $\nu_\infty^{\mathbb{N}}(\mathcal{E}^c) < \delta$  and

$$\frac{1}{n} \log \frac{\|g_{n-1}^\perp \cdots g_0^\perp v^\perp\|}{\|v^\perp\|} \in (\lambda_1(\nu_\infty) - \varepsilon, \lambda_1(\nu_\infty) + \varepsilon)$$

for any  $\mathbf{g} \in \mathcal{E}$  and  $n \geq N$ .

(iii)  $\eta_\infty(E) > 0$ .

We are going to see that the properties (i) - (iii) are incompatible with the fact that  $\eta_\infty$  is the limit of stationary measures for nearby random walks. Indeed, if  $E$  satisfies (i) and (ii) then it is a kind of repeller for the random walk on  $P$  associated to  $\nu_\infty$  (the precise statements are in Section 5.2). As we are going to see, that implies that the  $\eta_k$ -measure of any neighborhood of  $E$  is small when  $k$  is large, so that the limit  $\eta_\infty$  cannot satisfy (iii); the proof will use some general tools that we introduce in Sections 6.1 through 6.4. Hence the second alternative above cannot actually occur, and thus Theorem B will follow.

**5.1. Generic measures.** Actually, it suffices to carry these arguments in a special case, as we are going to explain. For each  $1 \leq r \leq d$ , let  $\text{Gr}(r, d)$  be the Grassmannian manifold of  $r$ -dimensional subspaces of  $\mathbb{R}^d$ . Moreover, let  $\mathcal{F}(r, d)$  be the space of flags

$$F_1 \subset F_2 \subset \cdots \subset F_{r-1} \subset F_r \subset \mathbb{R}^d,$$

where each  $F_i$  has dimension  $i$ . Note that  $\text{Gr}(1, d) = \mathcal{F}(1, d)$  coincides with the projective space  $P$ .

The natural action of  $G$  on the projective space extends to group actions on every  $\text{Gr}(r, d)$  and  $\mathcal{F}(r, d)$ . Thus, in particular, to each probability measure  $\nu$  on  $G$  we may associate operators  $\mathcal{P}_\nu$  and  $\mathcal{P}_\nu^*$  acting, respectively, on bounded measurable functions and on measures of  $\text{Gr}(r, d)$  or  $\mathcal{F}(r, d)$ , just as we did for  $P$  in (3.1) and (3.2):

$$(5.1) \quad \mathcal{P}_\nu \psi(v) = \int_G \psi(gv) d\nu(g) \text{ and } \mathcal{P}_\nu^* \eta = \int_G (g_* \eta) d\nu(g).$$

We continue to say that a function  $\psi$  is  $\nu$ -stationary if  $\mathcal{P}_\nu \psi = \psi$  and a probability measure  $\zeta$  is  $\nu$ -stationary if  $\mathcal{P}_\nu^* \zeta = \zeta$ .

A subset of an algebraic variety  $X$  is *Zariski-dense* if it is not contained in any proper algebraic subvariety of  $X$ . The cases we are interested in are  $X = \text{Gr}(r, d)$ ,  $\mathcal{F}(r, d)$ , or  $G$ . We call a measure  $\eta$  on the algebraic variety *generic* if  $\eta(M) = 0$  for any proper algebraic subvariety  $M$ . Then, in particular,  $\eta$  is *non-atomic*, meaning that  $\eta(\{p\}) = 0$  for every point  $p$  in the domain.

**Remark 5.1.** The restrictions  $\eta|_U$  of generic measures are generic, and so are the products  $\eta_1 \times \eta_2$  of generic measures. Moreover, if  $\{\eta_t : t \in T\}$  is a family of generic measures, and  $\xi$  is a measure on  $T$  then  $\eta = \int_T \eta_t d\xi(t)$  is a generic measure. In particular, the generic measures form a vector subspace.

**Proposition 5.2.** *Let  $X = \text{Gr}(r, d)$  or  $X = \mathcal{F}(r, d)$ . If  $\nu$  is a probability measure whose support is Zariski-dense in  $G$  then  $\nu$  admits a unique stationary measure  $\eta$  on  $X$ , and this measure is generic.*

*Proof.* Lemmas 4.2 and 4.5 in Gol'dsheid, Margulis [45]. □

**Proposition 5.3.** *Every  $\nu \in \mathcal{P}_c(G)$  is approximated by generic probability measures whose supports are Zariski-dense.*

*Proof.* For each  $j \in \mathbb{N}$ , let  $\xi_j \in \mathcal{P}_c(G)$  be the normalized restriction of the Haar measure of  $G$  to the ball of radius  $1/j$  around the identity  $I$  relative to some left-invariant distance on  $G$ . It is clear that  $\text{supp } \xi_j$  coincides with that ball, and  $\xi_j$  vanishes on any proper subvariety of  $G$ . Moreover,  $(\xi_j)_j$  converges to the Dirac mass  $\delta_I$  in the topology of  $\mathcal{P}_c(G)$ . For each  $j \in \mathbb{N}$ , let  $\nu_j = \nu * \xi_j$  be the probability measure on  $G$  such that, for any limited function  $\phi : G \rightarrow \mathbb{R}$ ,

$$\int_G \phi d\nu_j = \int_G \phi(gh) d\nu(g) d\xi_j(h).$$

The assumption on  $\xi_j$  ensures that  $\nu_j$  is generic. Moreover,  $(\nu_j)_j$  converges to  $\nu$  in the weak\* topology. Furthermore,  $\text{supp } \nu_j$  coincides with the  $(1/j)$ -neighborhood of the support of  $\nu$ , and so  $(\text{supp } \nu_j)_j \rightarrow \text{supp } \nu$  in the Hausdorff topology. This proves that  $(\nu_j)_j \rightarrow \nu$  in  $\mathcal{P}_c(G)$ . The fact that  $\text{supp } \nu_j$  has non-empty interior implies that it is Zariski-dense, for every  $j \in \mathbb{N}$ . □

**Corollary 5.4.** *Suppose that  $(\lambda_1(\nu'_k))_k \rightarrow \lambda_1(\nu_\infty)$  for any sequence  $(\nu'_k)_k$  converging to  $\nu_\infty$  in  $\mathcal{P}_c(G)$  such that  $\nu'_k$  is generic and  $\text{supp } \nu'_k$  is Zariski-dense in  $G$  for every  $k$ . Then  $(\lambda_1(\nu_k))_k \rightarrow \lambda_1(\nu_\infty)$  for every sequence  $(\nu_k)_k$  converging to  $\nu_\infty$  in  $\mathcal{P}_c(G)$ .*

*Proof.* Let  $(\nu_k)_k$  be any sequence converging to  $\nu_\infty$  in  $\mathcal{P}_c(G)$ , and  $d$  be any distance generating the topology of  $\mathcal{P}_c(G)$ . By Proposition 5.3, for each  $k$  we may find a generic probability measure  $\nu'_k$  arbitrarily close to  $\nu_k$  and whose support is Zariski-dense. Take  $\nu'_k$  such that  $d(\nu_k, \nu'_k) < 1/k$ . Using the well-known fact that the function  $\zeta \mapsto \lambda_1(\zeta)$  is upper semi-continuous (this is a consequence of (3.3) below), we may also suppose that  $\lambda_1(\nu'_k) \leq \lambda_1(\nu_k) + 1/k$ . Then  $(\nu'_k)_k \rightarrow \nu_\infty$  and so  $\lim_k \lambda_1(\nu'_k) = \lambda_1(\nu_\infty)$ . Moreover,

$$\liminf_k \lambda_1(\nu_k) \geq \lim_k \lambda_1(\nu'_k) = \lambda_1(\nu_\infty).$$

Using semi-continuity once more, this implies  $\lim_k \lambda_1(\nu_k) = \lambda_1(\nu_\infty)$ .  $\square$

Thus, to prove Theorem B it suffices to consider sequences  $(\nu_k)_k \rightarrow \nu_\infty$  of generic measures such that every  $\text{supp } \nu_k$  is Zariski-dense in  $G$ . We do so in all that follows. Then, by Proposition 5.2, the  $\nu_k$ -stationary measure  $\eta_k$  is unique and generic.

**5.2. Repelling behavior.** Given any subspace  $W \subset \mathbb{R}^d$ , let  $\Pi_W : \mathbb{R}^d \rightarrow W^\perp$  be the orthogonal projection along  $W$ . When  $W$  is 1-dimensional we also write  $\Pi_W = \Pi_w$  for any non-zero vector  $w \in W$ .

For any  $g \in G$  and  $v \in P$ , let  $Dg_v : T_v P \rightarrow T_{gv} P$  denote the derivative of  $g : G \rightarrow G$  at the point  $v \in P$ . The tangent space  $T_v P$  is naturally identified with the orthogonal space  $\{v\}^\perp$ . Then

$$(5.2) \quad Dg_v \dot{v} = \Pi_{gv} g \dot{v} \frac{\|v\|}{\|gv\|}.$$

For  $v \in E$ , let

$$(5.3) \quad Dg_v^\perp = \Pi_E \circ Dg_v | E^\perp : E^\perp \rightarrow E^\perp, \quad Dg_v^\perp v^\perp = g^\perp v^\perp \frac{\|v\|}{\|gv\|}.$$

For  $n \in \mathbb{N}$  and any probability measure  $\nu$  on  $G$ , let  $\nu^{(n)}$  denote the  $n$ -convolution, that is, the push-forward of  $\nu^\mathbb{N}$  under the map  $G^\mathbb{N} \rightarrow G$  defined by  $\mathbf{g} \mapsto g_{n-1} \cdots g_0$ :

$$(5.4) \quad \nu^{(n)} = \nu^\mathbb{N}(\{\mathbf{g} \in G^\mathbb{N} : g_{n-1} \cdots g_0 \in B\})$$

for any measurable set  $B \subset G$ . Note that the map  $\nu \mapsto \nu^{(n)}$  is continuous relative to the weak\* topology.

**Proposition 5.5.** *There exists  $\kappa_0 = \kappa_0(\nu_\infty) > 0$  and for each  $\delta > 0$  there exist  $N_0 = N_0(\nu_\infty, \delta) \in \mathbb{N}$  and  $\tau_0 = \tau_0(\nu_\infty, \delta) > 0$  such that for every  $n \geq N$  and  $v^\perp \in \mathbb{P}(E^\perp)$  there exists  $\mathcal{E}_0 = \mathcal{E}_0(\nu_\infty, \delta, n, v^\perp) \subset \text{supp } \nu_\infty^{(n)}$  with  $\nu_\infty^{(n)}(\mathcal{E}_0^c) < \delta$  and*

- (1)  $\log \frac{\|Dg_v^\perp v^\perp\|}{\|v^\perp\|} > \kappa_0 n$  for any  $g \in \mathcal{E}_0$  and  $v \in E$ ;
- (2)  $\frac{\|g^\perp v^\perp\|}{\|gv^\perp\|} > \tau_0$  for any  $g \in \mathcal{E}_0$ .

*Proof.* Fix numbers  $\alpha = \alpha(\nu_\infty)$  and  $\beta = \beta(\nu_\infty)$  such that

$$\lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 | E\| < \alpha < \beta < \lambda_1(\nu_\infty)$$

and then choose  $0 < \kappa_0 < \beta - \alpha$ . Let  $\delta > 0$ . By property (i), there exist  $N' \in \mathbb{N}$  and  $\mathcal{E}' \subset G^{\mathbb{N}}$  such that  $\nu_{\mathcal{X}}^{\mathbb{N}}((\mathcal{E}')^c) < \delta/4$  and

$$(5.5) \quad \log \frac{\|g_{n-1} \cdots g_0 v\|}{\|v\|} \leq \alpha n \text{ for any } \mathbf{g} \in \mathcal{E}' \text{ and } v \in E \text{ and } n \geq N'.$$

By property (ii), there exists  $N'' \in \mathbb{N}$  and for each  $v^\perp \in E^\perp$  there exists  $\mathcal{E}'' \subset G^{\mathbb{N}}$  such that  $\nu_{\mathcal{X}}^{\mathbb{N}}((\mathcal{E}'')^c) < \delta/4$  and

$$(5.6) \quad \log \frac{\|g_{n-1}^\perp \cdots g_0^\perp v^\perp\|}{\|v^\perp\|} \geq \beta n \text{ for any } \mathbf{g} \in \mathcal{E}'' \text{ and } n \geq N''.$$

Let  $N''' = \max\{N', N''\}$  and for each  $n \geq N'''$  and  $v^\perp \in \mathbb{P}(E^\perp)$  define

$$\mathcal{E}''' = \{g_{n-1} \cdots g_0 : \mathbf{g} \in \mathcal{E}' \cap \mathcal{E}''\}.$$

The definition (5.4) gives that  $\nu_{\mathcal{X}}^{(n)}((\mathcal{E}''')^c) \leq \nu_{\mathcal{X}}^{\mathbb{N}}((\mathcal{E}')^c \cup (\mathcal{E}'')^c) < \delta/2$ . Moreover, if  $\mathbf{g} \in \mathcal{E}'''$  then (5.5) and (5.6) give that

$$(5.7) \quad \log \frac{\|Dg_v^\perp v^\perp\|}{\|v^\perp\|} = \log \frac{\|g^\perp v^\perp\|}{\|v^\perp\|} - \log \frac{\|gv\|}{\|v\|} \geq (\beta - \alpha)n > \kappa_0 n$$

for any  $v \in E$  and  $n \geq N'''$ . This gives claim (1), as long as we choose  $N_0 \geq N'''$  and  $\mathcal{E}_0 \subset \mathcal{E}'''$ , which we will do in the next paragraph.

Now let us explain how to obtain claim (2). The following elementary inequality will be used a couple of times:

$$(5.8) \quad |\sin \angle(v_1, v_2)| \leq \frac{\|v_1 \pm v_2\|}{\|v_2\|} \text{ for any non-zero } v_1, v_2 \in \mathbb{R}^d.$$

Let  $a = (\chi_1 - \chi_2)/8$  where  $\chi_1 > \chi_2$  are the two largest Lyapunov exponents in (2.1). Clearly, we may assume that  $\beta$  has been chosen greater than  $\chi_1 - a$  (keep in mind that  $\chi_1 = \lambda_1$ ). Consider the Oseledets splitting in (2.3) and denote  $E^* = E^2 \oplus \cdots \oplus E^k$ . By the Oseledets theorem (see [65, Theorem 4.2]), there exists a measurable function  $\mathbf{g}_\pm \mapsto c(\mathbf{g}_\pm)$  with values in  $(0, 1)$  such that

$$(5.9) \quad c(\mathbf{g}_\pm) e^{n(\chi_1 - a)} \|v_1\| \leq \|g_{n-1} \cdots g_0 v_1\| \leq c(\mathbf{g}_\pm)^{-1} e^{n(\chi_1 + a)} \|v_1\|$$

$$(5.10) \quad \|g_{n-1} \cdots g_0 v_*\| \leq c(\mathbf{g}_\pm)^{-1} e^{n(\chi_2 + a)} \|v_*\|$$

for any  $v_1 \in E^1(\mathbf{g}_\pm)$ ,  $v_* \in E^*(\mathbf{g}_\pm)$  and  $n \in \mathbb{N}$ , and for  $\nu_{\mathcal{X}}^{\mathbb{Z}}$ -almost every  $\mathbf{g}_\pm$ . In particular,  $E^1(\mathbf{g}_\pm) \cap E^*(\mathbf{g}_\pm) = \{0\}$ . Since the growth rate of every vector  $v \in E$  is strictly less than  $\lambda_1 = \chi_1$  (property (i) above), we also have that  $E^1(\mathbf{g}_\pm) \cap E = \{0\}$ . Thus, up to reducing the function  $c(\mathbf{g}_\pm)$ , we may suppose that

$$(5.11) \quad |\sin \angle(E^1(\mathbf{g}_\pm), E^*(\mathbf{g}_\pm))| \geq c(\mathbf{g}_\pm) \text{ and } |\sin \angle(E^1(\mathbf{g}_\pm), E)| \geq c(\mathbf{g}_\pm)$$

for  $\nu_{\mathcal{X}}^{\mathbb{Z}}$ -almost every  $\mathbf{g}_\pm$ . Fix  $b = b(\nu_{\mathcal{X}}, \delta) > 0$  small enough that the set

$$\mathcal{A} = \{\mathbf{g}_\pm \in G^{\mathbb{Z}} : c(\mathbf{g}_\pm) \geq b\}$$

has  $\nu_{\mathcal{X}}^{\mathbb{Z}}(\mathcal{A}^c) < \delta/4$ . Then fix  $M = M(\nu_{\mathcal{X}}, \delta) \in \mathbb{N}$  such that

$$(5.12) \quad e^{-an} < \frac{b^2}{2} \text{ for every } n \geq M.$$

Let  $N_0 = \max\{N', N'', M\}$  and for each  $n \geq N_0$  and  $v^\perp \in \mathbb{P}(E^\perp)$  define

$$\mathcal{E}_0 = \{g_{n-1} \cdots g_0 : \mathbf{g}_\pm \in \pi^{-1} \mathcal{E}' \cap \pi^{-1} \mathcal{E}'' \cap \mathcal{A} \cap \sigma^{-n} \mathcal{A}\} \cap \text{supp } \nu_{\mathcal{X}}^{(n)}$$

( $\pi : G^{\mathbb{Z}} \rightarrow G^{\mathbb{N}}$  is the canonical projection). The choices of  $\mathcal{E}'$ ,  $\mathcal{E}''$  and  $\mathcal{A}$  ensure that  $\nu_{\infty}^{(n)}(\mathcal{E}_0^c) < \delta$ . For each  $g \in \mathcal{E}_0$ , take  $\mathfrak{g}_{\pm} \in \mathcal{E}' \cap \mathcal{E}'' \cap \mathcal{A} \cap \sigma^{-n}\mathcal{A}$  such that  $g = g_{n-1} \cdots g_0$  and then let  $v^{\perp} = v_1 + v_*$  be the decomposition of  $v^{\perp}$  with respect to the splitting  $\mathbb{R}^d = E^1(\mathfrak{g}_{\pm}) \oplus E^*(\mathfrak{g}_{\pm})$ . From property (5.11) and the fact that  $\mathfrak{g}_{\pm} \in \mathcal{A}$  we get

$$(5.13) \quad \|v_*\| \leq \frac{\|v^{\perp}\|}{|\sin \angle(E^1(\mathfrak{g}_{\pm}), E^*(\mathfrak{g}_{\pm}))|} \leq c(\mathfrak{g}_{\pm})^{-1} \|v^{\perp}\| \leq b^{-1} \|v^{\perp}\|.$$

Since  $\beta > \chi_1 - a$ , the inequality (5.6) gives

$$(5.14) \quad \|gv^{\perp}\| \geq \|g^{\perp}v^{\perp}\| \geq e^{(\chi_1 - a)n} \|v^{\perp}\|.$$

Properties (5.9) and (5.10) give

$$(5.15) \quad \|gv^{\perp}\| \leq b^{-1} e^{(\chi_1 + a)n} \|v_1\| + b^{-1} e^{(\chi_2 + a)n} \|v_*\|.$$

Putting the relations (5.13) through (5.15) together, and using (5.12), we obtain

$$\begin{aligned} \|v^{\perp}\| &\leq b^{-1} e^{2an} \|v_1\| + b^{-1} e^{(\chi_2 - \chi_1 + 2a)n} \|v_*\| \\ &\leq b^{-1} e^{2an} \|v_1\| + b^{-2} e^{-6an} \|v^{\perp}\| \leq b^{-1} e^{2an} \|v_1\| + \frac{1}{2} \|v^{\perp}\| \end{aligned}$$

This proves that  $\|v_1\| \geq (b/2) e^{-2an} \|v^{\perp}\|$ . Combining this inequality with (5.13) and properties (5.9) and (5.10),

$$\begin{aligned} \|gv_1\| &\geq b e^{(\chi_1 - a)n} \|v_1\| \geq (b^2/2) e^{(\chi_1 - 3a)n} \|v^{\perp}\| \\ \|gv_*\| &\leq b^{-1} e^{(\chi_2 + a)n} \|v_*\| \leq b^{-2} e^{(\chi_2 + a)n} \|v^{\perp}\| \end{aligned}$$

In view of our choice of  $a$  and the relation (5.13), this implies that

$$\frac{\|gv_*\|}{\|gv_1\|} \leq 2b^{-4} e^{(\chi_2 - \chi_1 + 4a)n} \leq 2b^{-4} e^{-4an} < \frac{b}{2} e^{-2an} \leq \frac{b}{2}.$$

Then it follows that

$$|\sin \angle(gv^{\perp}, E^1(\sigma^n \mathfrak{g}_{\pm}))| \leq |\sin \angle(gv^{\perp}, gv_1)| \leq \frac{\|gv_*\|}{\|gv_1\|} < b/2.$$

Now, property (5.11) implies that  $|\sin \angle(E^1(\sigma^n \mathfrak{g}_{\pm}), E)| \geq b$ , because we have taken  $\mathfrak{g}_{\pm}$  such that  $\sigma^n \mathfrak{g}_{\pm} \in \mathcal{A}$ . Since  $|\sin|$  is a subadditive function, these two inequalities imply that  $|\sin \angle(gv^{\perp}, E)| > b/2$ , which means that  $\|g^{\perp}v^{\perp}\| > (b/2) \|gv^{\perp}\|$ . This proves (2) with  $\tau_0 = b/2$ .  $\square$

Let  $d(\cdot, \cdot)$  be the distance defined on the projective space  $P$  by

$$(5.16) \quad d(u, v) = \left\| \Pi_u \frac{v}{\|v\|} \right\| = |\sin \angle(u, v)|.$$

Note that  $0 \leq d(u, v) \leq 1$  for every  $u, v \in P$ . Next, we formulate the infinitesimal estimate in part (1) of Proposition 5.5 in terms of the distance to the equator:

**Corollary 5.6.** *For each  $n \geq N$  there exists  $\rho_0 = \rho_0(\nu_{\infty}, n) > 0$  and for each  $x \in P$  with  $d(x, E) \leq \rho_0$  there exists  $\mathcal{D}_{\infty}(x) = \mathcal{D}_{\infty}(\nu_{\infty}, \delta, n, x) \subset \text{supp } \nu_{\infty}^{(n)}$  with  $\nu_{\infty}^{(n)}(\mathcal{D}_{\infty}(x)^c) < \delta$  and*

$$(5.17) \quad -\log d(gx, E) \leq -\log d(x, E) - \frac{3\kappa_0}{4} n \quad \text{for any } g \in \mathcal{D}_{\infty}(x).$$



*Proof.* Let  $\exp$  denote the exponential map of the Riemannian manifold  $P$ . For each  $x \in P$  close to the equator we may write  $x = \exp_v v^\perp$  for a (unique)  $v^\perp \in E^\perp$  with  $\|v^\perp\| = d(x, v) = d(x, E)$ . Then

$$\lim_{x \rightarrow E} \frac{d(gx, E)}{\|Dg_v^\perp v^\perp\|} = \lim_{x \rightarrow E} \frac{d(\exp_{gv} Dg_v v^\perp, E)}{\|Dg_v^\perp v^\perp\|} = 1$$

and the limits are uniform in  $g \in \text{supp } \nu_\infty^{(n)}$ . In particular, there exists  $\rho_0 = \rho_0(\nu_\infty, n) > 0$  such that

$$(5.18) \quad d(x, E) \leq \rho_0 \quad \Rightarrow \quad \left| \log \frac{d(gx, E)}{d(x, E)} - \log \frac{\|Dg_v^\perp v^\perp\|}{\|v^\perp\|} \right| < \frac{\kappa_0}{4} n$$

for every  $g \in \text{supp } \nu_\infty^{(n)}$ . Define  $\mathcal{D}_\infty(x)$  to be the set  $\mathcal{E}_0(\nu_\infty, \delta, n, v^\perp)$  given by Proposition 5.5. Then (5.17) follows from (5.18) and part (1) of Proposition 5.5.  $\square$

We also need to extend these estimates from  $k = \infty$  to every large  $k \in \mathbb{N}$ :

**Corollary 5.7.** *For any  $n \geq N$  and  $\rho \in (0, \rho_0)$  there is  $k_0 = k_0(\nu_\infty, \delta, n, \rho) \in \mathbb{N}$  and for any  $x \in P$  with  $\rho \leq d(x, E) \leq \rho_0$  there is  $\mathcal{D}_k(x) = \mathcal{D}_k(\nu_\infty, \delta, n, \rho, x) \subset \text{supp } \nu_k^{(n)}$  such that  $\nu_k^{(n)}(\mathcal{D}_k(x)^c) < \delta$  and*

$$(5.19) \quad -\log d(gx, E) \leq -\log d(x, E) - \frac{\kappa_0}{2} n \quad \text{for any } g \in \mathcal{D}_k(x) \text{ and } k \geq k_0.$$

*Proof.* Let  $n \geq N$  and  $\rho \in (0, \rho_0)$  be fixed. The set  $K = \{x \in P : \rho \leq d(x, E) \leq \rho_0\}$  is compact. By continuity, there exists  $\theta = \theta(\nu_\infty, n, \rho) > 0$  (keep in mind that  $\kappa_0 = \kappa_0(\nu_\infty)$  and  $\rho_0 = \rho_0(\nu_\infty, n)$ ), such that

$$(5.20) \quad \left| \log \frac{d(gx, E)}{d(x, E)} - \log \frac{d(hy, E)}{d(y, E)} \right| < \frac{\kappa_0}{4} n$$

for any  $g \in B(h, \theta)$  and  $x \in B(y, \theta)$  and  $h \in \text{supp } \nu_\infty^{(n)}$  and  $y \in K$ . Choose a finite set  $\{x_j : j = 1, \dots, l\} \subset K$  such that the balls  $B(x_j, \theta)$ ,  $j = 1, \dots, l$  cover  $K$ . For each  $x \in K$  choose  $j \in \{1, \dots, l\}$  such that  $x \in B(x_j, \theta)$  and define

$$(5.21) \quad \mathcal{D}_k(x) = [\theta\text{-neighborhood of } \mathcal{D}_\infty(x_j)] \cap \text{supp } \nu_k^{(n)}.$$

Let  $x \in K$  and  $g \in \mathcal{D}_k(x)$ . By definition, there exist  $x_j \in K$  and  $h \in \mathcal{D}_\infty(x_j)$  such that  $d(x, x_j) < \theta$  and  $d(g, h) < \theta$ . Then (5.20) gives that

$$-\log d(gx, E) \leq -\log d(x, E) - \log d(hx_j, E) + \log d(x_j, E) + \frac{\kappa_0}{4} n,$$

whereas (5.17) gives that  $-\log d(hx_j, E) \leq -\log d(x_j, E) - (3/4)\kappa_0 n$ . Substituting the latter in the former one obtains (5.19). Since  $\nu_k^{(n)}$  converges to  $\nu_\infty^{(n)}$  in the weak\* topology, the limit inferior of the  $\nu_k^{(n)}$ -measures of (5.21) is greater than or equal to

$$\nu_\infty^{(n)}(\mathcal{D}_\infty(x_j)) > 1 - \delta$$

for any  $j = 1, \dots, l$ . In particular, there is  $k_0 = k_0(\nu_\infty, \delta, n, \rho) \in \mathbb{N}$  such that

$$\nu_k^{(n)}(\mathcal{D}_k(x)) > 1 - \delta \text{ for every } k \geq k_0 \text{ and } x \in K.$$

This completes the argument.  $\square$

## 6. A TOOLBOX

Here we collect several fairly general ideas and facts that are required for the continuation of our arguments. The proofs themselves will not be used in what follows, so the reader is advised to skip them at this stage, and to return to this section for reference along the way, as needed. In our applications, the metric spaces  $X$  and  $X'$  will be suitable subsets of Grassmannian manifolds or flag varieties.

**6.1. Couplings.** Let  $\eta$  and  $\eta'$  be measures on measurable spaces  $X$  and  $X'$ , respectively, with  $\eta(X) = \eta'(X')$ . A measure  $\tilde{\eta}$  on  $X \times X'$  is a *coupling* of  $\eta$  and  $\eta'$  if it projects to  $\eta$  on the first factor and to  $\eta'$  on the second factor, that is, if

$$\tilde{\eta}(A \times X') = \eta(A) \text{ and } \tilde{\eta}(X \times A') = \eta'(A')$$

for any measurable sets  $A \subset X$  and  $A' \subset X'$ . When  $\eta = \eta'$  (and  $X = X'$ ) we call  $\tilde{\eta}$  a *self-coupling* of  $\eta$ . A self-coupling is *symmetric* if it is invariant under the involution  $\iota : (x, x') \mapsto (x', x)$ .

For example, the *diagonal embedding* of a measure  $\eta$  on  $X$  is the symmetric self-coupling  $\tilde{\eta}$  of  $\eta$  defined by

$$(6.1) \quad \tilde{\eta}(B) = \eta(\{x \in X : (x, x) \in B\}).$$

Another simple example of a coupling is the *scaled product*

$$(6.2) \quad \tilde{\eta} = \frac{1}{c}(\eta \times \eta'), \text{ where } c = \eta(X) = \eta'(X').$$

Couplings are far from being unique, which turns out to be a very convenient feature in our context. Especially, we will use the following elementary construction:

**Lemma 6.1.** *Suppose that  $A \subset X$  and  $A' \subset X'$  are such that  $\eta(A) < \eta(X \setminus A)$  and  $\eta'(A') < \eta'(X' \setminus A')$ . Then there exists a coupling  $\tilde{\eta}$  of  $\eta$  and  $\eta'$  such that  $\tilde{\eta}(A \times A') = 0$ .*

*Proof.* Let  $B = X \setminus A$  and  $B' = X' \setminus A'$  and then take

$$(6.3) \quad \begin{aligned} \tilde{\eta} = & \frac{1}{\eta'(B')}(\eta|_A \times (\eta'|_{B'})) + \frac{1}{\eta(B)}(\eta|_B \times (\eta'|_{A'})) \\ & + \left( \frac{1}{\eta(B)} + \frac{1}{\eta'(B')} - \frac{c}{\eta(B)\eta'(B')} \right) (\eta|_B \times (\eta'|_{B'}), \end{aligned}$$

where  $c = \eta(X) = \eta'(X')$ . □

**Lemma 6.2.** *Let  $A_j \subset C_j \subset X$  and  $A'_j \subset C'_j \subset X'$ ,  $j = 1, \dots, k$  be such that*

- (1)  $\eta(C_j) < \eta(X \setminus C_j)$  and  $\eta'(C'_j) < \eta'(X' \setminus C'_j)$  for  $1 \leq j \leq k$ ;
- (2)  $A_j \times (X' \setminus C'_j)$  and  $(X \setminus C_j) \times A'_j$  are disjoint from  $A_i \times A'_i$  for  $1 \leq i < j \leq k$ .

*Then there exists a coupling  $\tilde{\eta}$  of  $\eta$  and  $\eta'$  such that  $\tilde{\eta}(A_j \times A'_j) = 0$  for  $j = 1, \dots, k$ .*

*Proof.* We are going to construct couplings  $\tilde{\eta}_j$ ,  $j = 1, \dots, k$  of  $\eta$  and  $\eta'$  such that

$$(6.4) \quad \tilde{\eta}_j(A_i \times A'_i) = 0 \text{ for any } 1 \leq i \leq j.$$

Then it suffices to take  $\tilde{\eta} = \tilde{\eta}_k$ .

The case  $j = 1$  of (6.4) is contained in Lemma 6.1. We proceed by induction. Let  $j \in \{2, \dots, k\}$  and assume that we have constructed a coupling  $\tilde{\eta}_{j-1}$  of  $\eta$  and  $\eta'$  such that  $A_i \times A'_i$  has zero measure for every  $i = 1, \dots, j-1$ . We claim that

$$(6.5) \quad \tilde{\eta}_{j-1}(C_j \times C'_j) < \tilde{\eta}_{j-1}(C_j^c \times (C'_j)^c),$$

where  $C_j^c = X \setminus C_j$  and  $(C_j')^c = X' \setminus C_j'$ . Indeed, suppose that  $\eta'(C_j') \leq \eta(C_j)$ . Recall that  $\eta(C_j) < \eta(C_j^c)$ , by assumption. Moreover,

- $\tilde{\eta}_{j-1}(C_j \times C_j') + \tilde{\eta}_{j-1}(C_j^c \times C_j') = \eta'(C_j')$ ,
- $\tilde{\eta}_{j-1}(C_j^c \times C_j') + \tilde{\eta}_{j-1}(C_j^c \times (C_j')^c) = \eta(C_j^c)$ .

Combining these relations we get the inequality in (6.5). The case  $\eta(C_j) \leq \eta'(C_j')$  is analogous, and so the claim is proved. In particular, the following number is smaller than 1:

$$\theta_j = \frac{\tilde{\eta}_{j-1}(A_j \times A_j')}{\tilde{\eta}_{j-1}(C_j^c \times (C_j')^c)}.$$

Let  $\pi : X \times X' \rightarrow X$  and  $\pi' : X \times X' \rightarrow X'$  be the canonical projections. Let  $\zeta_j$  be the scaled product of  $\pi_*(\tilde{\eta}_{j-1} | A_j \times A_j')$  and  $\theta_j \pi'_*(\tilde{\eta}_{j-1} | C_j^c \times (C_j')^c)$ , and  $\zeta_j'$  be the scaled product of  $\theta_j \pi_*(\tilde{\eta}_{j-1} | C_j^c \times (C_j')^c)$  and  $\pi'_*(\tilde{\eta}_{j-1} | A_j \times A_j')$ . Then define

$$(6.6) \quad \tilde{\eta}_j = \tilde{\eta}_{j-1} - (\tilde{\eta}_{j-1} | A_j \times A_j') - \theta_j (\tilde{\eta}_{j-1} | C_j^c \times (C_j')^c) + \zeta_j + \zeta_j'.$$

It is clear that  $\tilde{\eta}_j$  is a (positive) measure, because  $\theta_j < 1$ . It is also clear that  $\tilde{\eta}_j$  is a coupling of  $\eta$  and  $\eta'$ :

$$\begin{aligned} \pi_* \tilde{\eta}_j &= \pi_* \tilde{\eta}_{j-1} - \pi_*(\tilde{\eta}_{j-1} | A_j \times A_j') - \theta_j \pi_*(\tilde{\eta}_{j-1} | C_j^c \times (C_j')^c) \\ &\quad + \pi_*(\tilde{\eta}_{j-1} | A_j \times A_j') + \theta_j \pi_*(\tilde{\eta}_{j-1} | C_j^c \times (C_j')^c) = \pi_* \tilde{\eta}_{j-1} = \eta \end{aligned}$$

and, analogously,  $\pi'_* \tilde{\eta}_j = \eta'$ . Note also that  $\tilde{\eta}_j(A_j \times A_j') = 0$ . Moreover,  $\zeta_j$  is concentrated in  $A_j \times (C_j')^c$  and  $\zeta_j'$  is concentrated in  $C_j^c \times A_j'$ . Since both sets are assumed to be disjoint from  $A_i \times A_i'$ , we get that

$$\tilde{\eta}_j(A_i \times A_i') \leq \tilde{\eta}_{j-1}(A_i \times A_i') = 0 \text{ for } i = 1, \dots, j-1.$$

This completes the induction.  $\square$

**Remark 6.3.** By construction, the coupling  $\tilde{\eta}$  varies continuously with  $\eta$  and  $\eta'$  in the weak\*-topology at all points such that the boundaries of all  $A_j$  and  $C_j$  have zero  $\eta$ -measure and the boundaries of all  $A_j'$  and  $C_j'$  have zero  $\eta'$ -measure.

**Remark 6.4.** The constructions in Lemmas 6.1 and 6.2 preserve the class of non-atomic measures and, when  $X$  is an algebraic variety, also the class of generic measures: if  $\eta$  and  $\eta'$  are generic then so is the coupling  $\tilde{\eta}$ . That follows directly from Remark 5.1 and the expressions (6.3) and (6.6).

**Remark 6.5.** When  $X = X'$ , these constructions are *involution-invariant* in the following sense. First, in Lemma 6.1 the coupling  $\tilde{\eta}$  given by (6.3) is replaced with  $\iota_* \tilde{\eta}$  when one exchanges the roles of  $\eta, A$  and  $\eta', A'$ . In particular, if  $\eta = \eta'$  and  $A = A'$  then the self-coupling  $\tilde{\eta}$  is symmetric. In Lemma 6.2 we may take  $\zeta_j' = \iota_* \zeta_j$  and then the coupling  $\tilde{\eta}_j$  is replaced with  $\iota_* \tilde{\eta}_j$  when the roles of  $\eta, A_i, C_i$  and  $\eta', A_i', C_i'$  are exchanged. In particular, if  $\eta = \eta', A_i = A_i',$  and  $C_i = C_i'$  then the self-coupling  $\tilde{\eta} = \tilde{\eta}_k$  is symmetric.

**Proposition 6.6.** *Let  $K$  be a compact subset of the product  $X \times X'$  such that*

$$K(x') = \{y \in X : (y, x') \in K\} \text{ and } K'(x) = \{y' \in X' : (x, y') \in K\}.$$

*satisfy*

$$(6.7) \quad \eta(K(x')) < \eta(X \setminus K(x')) \text{ and } \eta'(K'(x)) < \eta'(X' \setminus K'(x))$$

for every  $(x, x') \in X \times X'$ . Then there exists a coupling  $\tilde{\eta}$  of  $\eta$  and  $\eta'$  that vanishes on a neighborhood of  $K$ .

*Proof.* We are going to find open sets  $A_j \subset C_j \subset X$  and  $A'_j \subset C'_j \subset X'$ ,  $j = 1, \dots, k$ , such that

- (a)  $\eta(C_j) < \eta(X \setminus C_j)$  and  $\eta(C'_j) < \eta(X \setminus C'_j)$ ;
- (b)  $A_j \times A'_j$  is disjoint from  $(X \setminus C_i) \times A'_i$  and  $A_i \times (X \setminus C'_i)$ ;
- (c) and the union of the products  $A_j \times A'_j$  contains  $K$ .

Then, by Lemma 6.2, there exists a coupling  $\tilde{\eta}$  of  $\eta$  and  $\eta'$  vanishing on the union of  $A_j \times A'_j$ ,  $j = 1, \dots, k$ , which gives the claim of the present lemma. So let us explain how to construct such sets.

By compactness, there exists  $\varepsilon > 0$  such that (6.7) remains valid when  $K(x')$  and  $K'(x)$  are replaced with their  $10\varepsilon$ -neighborhoods. Let  $\varepsilon$  be fixed. Also by compactness, the maps  $x \mapsto K'(x)$  and  $x' \mapsto K(x')$  are upper semicontinuous. In particular, for any  $x \in X$  and  $x' \in X'$  there exist  $\delta(x) > 0$  and  $\delta'(x') > 0$  such that

$$(6.8) \quad \begin{aligned} K'(y) &\subset B_\varepsilon(K'(x)) \text{ if } d(x, y) < 4\delta(x) \text{ and} \\ K(y') &\subset B_\varepsilon(K(x')) \text{ if } d(x', y') < 4\delta(x'). \end{aligned}$$

It is no restriction to assume that  $\delta(x)$  and  $\delta'(x')$  are bounded by  $\varepsilon$ , and we do so. Let  $(x_1, x'_1), \dots, (x_k, x'_k) \in K$  be such that  $B(x_j, \delta(x_j)) \times B(x'_j, \delta'(x'_j))$ ,  $j = 1, \dots, k$  cover  $K$ . Initially, define

$$(6.9) \quad \begin{aligned} A_j &= B(x_j, 3\delta(x_j)), & A'_j &= B(x'_j, 3\delta'(x'_j)) \\ C_j &= B(K(x'_j), 10\varepsilon), & C'_j &= B(K'(x_j), 10\varepsilon). \end{aligned}$$

Note that  $A_j \subset C_j$  and  $A'_j \subset C'_j$  since  $(x_j, x'_j) \in K$  and  $\delta, \delta' \leq \varepsilon$ . Conditions (a) and (c) are clearly satisfied, but not necessarily (b). In the following we replace the sets  $A_j$  and  $A'_j$  with suitable subsets, in such a way as to achieve (b) while preserving (c). The condition (a) is clearly not affected.

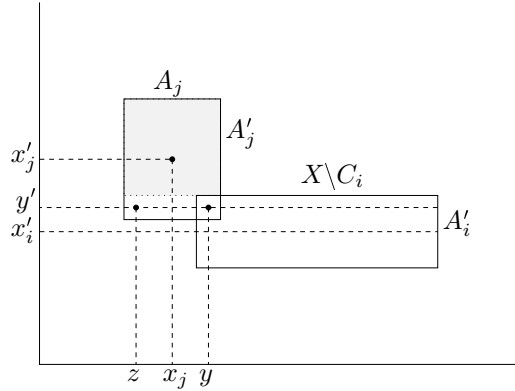


FIGURE 1. Trimming the sets  $A_j$  and  $A'_j$ : to achieve the disjointness condition (c), in the situation described in the figure  $A_j \times A'_j$  is replaced with  $A_j \times (A'_j \setminus \bar{A}'_i)$ , which corresponds to the shaded region. A dual operation is applied to  $A_j$ , and the whole procedure is repeated for every  $i$  and  $j$ .

Consider any  $i, j = 1, \dots, k$ . If  $A_j \times A'_j$  is disjoint from  $A_i \times (X' \setminus C'_i)$  and  $(X \setminus C_i) \times A'_i$  there is nothing to do. Next, let us consider the case when there exists  $(y, y')$  in  $(A_j \times A'_j) \cap ((X \setminus C_i) \times A'_i)$ . See Figure 1. We claim that  $A_j \times B(x'_i, 4\delta(x'_i))$  is disjoint from  $K$ . Indeed, suppose there existed  $z \in A_j$  and  $y'' \in B(x'_i, 4\delta(x'_i))$  such that  $(z, y'') \in K$ . Since  $d(z, y) < 6\delta(x_j) \leq 6\varepsilon$ , it would follow that  $y \in B(K(y''), 6\varepsilon)$ . On the other hand,  $K(y'') \subset B(K(x'_i), \varepsilon)$  because  $d(y'', x'_i) < 4\delta(x'_i)$ . Hence, we would have  $y \in B(K(x'_i), 7\varepsilon)$ , which would contradict the fact that  $y \in X \setminus C_i$ . This contradiction proves our claim. Now, this ensures that  $A_j \times \bar{A}'_i$  is disjoint and at a definite distance from  $K$ . Thus, we may replace  $A'_j$  with  $A'_j \setminus \bar{A}'_i$  in our construction without affecting condition (d) and, by doing it, we get that  $A_j \times A'_j$  becomes disjoint from  $(X \setminus C_i) \times A'_i$ . The case when there exists  $(y, y')$  in  $(A_j \times A'_j) \cap (A_i \times (X' \setminus C'_i))$  is treated in the same way, trimming  $A'_j$  instead. Repeating this procedure for each  $i$  and  $j$ , we get all three conditions (a) to (c).  $\square$

**Remark 6.7.** The union of the boundaries  $\partial A_j$  over all  $j = 1, \dots, k$  does not increase under the trimming operation, and the same holds for the union of the boundaries  $\partial A'_j$  over  $j = 1, \dots, k$ .

Next we want to state and prove a parametrized version of Proposition 6.6. The following elementary fact will be useful at other places as well:

**Lemma 6.8.** *Let  $g : Z \times T \rightarrow \mathbb{R}$  be a bounded measurable function, where  $Z$  is a metric space and  $(T, \mu)$  is a probability space. Let  $z_0 \in Z$  be such that the set  $D(z_0)$  of values of  $t \in T$  such that  $z \mapsto g(z, t)$  is discontinuous at  $z_0$  has zero  $\mu$ -measure. Then  $z \mapsto \int_T g(z, t) d\mu(t)$  is continuous at  $z = z_0$ .*

*Proof.* Fix any  $\varepsilon > 0$ . For each  $k \in \mathbb{N}$ , denote by  $T_k$  the set of values of  $t \in T$  such that  $|g(z_0, t) - g(z, t)| \leq \varepsilon$  for any  $z$  in the  $(1/k)$ -neighborhood of  $z_0$ . The sequence  $T_k$  is non-decreasing and the assumption ensures that  $\cup_k T_k$  has full  $\mu$ -measure. Fix  $k$  such that  $\mu(T_k^c) < \varepsilon$ . Then for any  $z$  in the  $(1/k)$ -neighborhood of  $z_0$ ,

$$\left| \int_T g(z_0, \cdot) d\mu - \int_T g(z, \cdot) d\mu \right| \leq \varepsilon + \left| \int_{T_k^c} g(z_0, \cdot) d\mu \right| + \left| \int_{T_k^c} g(z, \cdot) d\mu \right| \leq (1 + 2\|g\|)\varepsilon.$$

Since  $\varepsilon > 0$  is arbitrary, this proves that  $z_0$  is a continuity point.  $\square$

**Proposition 6.9.** *Let  $X, X', Y$ , and  $Y'$  be compact metric spaces,  $K$  be a compact subset of  $X \times X'$ , and  $\eta = \{\eta_y : y \in Y\}$  and  $\eta' = \{\eta'_{y'} : y' \in Y'\}$  be continuous families of probability measures on  $X$  and  $X'$ , respectively, such that*

$$(6.10) \quad \eta_y(K(x')) < \eta_y(X \setminus K(x')) \quad \text{and} \quad \eta'_{y'}(K'(x)) < \eta'_{y'}(X' \setminus K'(x)).$$

*for every  $(x, x') \in X \times X'$  and  $(y, y') \in Y \times Y'$ . Then there exists a continuous family  $\tilde{\eta} = \{\tilde{\eta}_{y, y'} : (y, y') \in Y \times Y'\}$  of probability measures on  $X \times X'$  such that each  $\tilde{\eta}_{y, y'}$  is a coupling of  $\eta_y$  and  $\eta'_{y'}$  vanishing on a uniform neighborhood of  $K$ .*

*Proof.* By compactness, the assumption (6.10) implies that there exists  $\varepsilon > 0$  such that

$$\begin{aligned} \eta_y(B(K(x'), 10\varepsilon)) &< \eta_y(X \setminus B(K(x'), 10\varepsilon)) \quad \text{and} \\ \eta'_{y'}(B(K'(x), 10\varepsilon)) &< \eta'_{y'}(X' \setminus B(K'(x), 10\varepsilon)) \end{aligned}$$

for every  $(x, x') \in X \times X'$  and  $(y, y') \in Y \times Y'$ . Fix  $\varepsilon$  and let  $\delta(x), \delta'(x') \in (0, \varepsilon)$  be as in (6.8). Let  $(x_1, x'_1), \dots, (x_k, x'_k) \in K$  be such that  $B(x_j, \delta(x_j)) \times B(x'_j, \delta'(x'_j))$ ,

$j = 1, \dots, k$  cover  $K$ . For each  $s \in [0, 1]$  and  $j = 1, \dots, k$ , define

$$(6.11) \quad \begin{aligned} A_{j,s} &= B(x_j, (3-s)\delta(x_j)), & A'_{j,s} &= B(x'_j, (3-s)\delta'(x'_j)) \\ C_{j,s} &= B(K(x'_j), (10-s)\varepsilon), & C'_{j,s} &= B(K'(x_j), (10-s)\varepsilon). \end{aligned}$$

Applying to the measures  $\eta_y$  and  $\eta'_{y'}$ , and the sets  $A_{j,s}$ ,  $C_{j,s}$ ,  $A'_{j,s}$ , and  $C'_{j,s}$  the same construction in Lemma 6.2, we find for each  $s \in [0, 1]$  and  $(y, y') \in Y \times Y'$  a coupling  $\tilde{\eta}_{y,y',s}$  of  $\eta_y$  and  $\eta'_{y'}$  which vanishes on a neighborhood of  $K$  independent of both  $s$  and  $(y, y')$ . (recall that the trimming is always done at a definite distance from  $K$ ).

By Remark 6.3, for each fixed  $s \in [0, 1]$  the map  $(y, y') \mapsto \tilde{\eta}_{y,y',s}$  is discontinuous at a given point  $(z, z') \in Y \times Y'$  only if the union of the boundaries of  $A_{j,s}$  and  $C_{j,s}$  has positive measure for  $\eta_z$  or the union of the boundaries of  $A'_{j,s}$  and  $C'_{j,s}$  has positive measure for  $\eta'_{z'}$ . In the setting of (6.11), the boundaries of the  $A_{j,s}$  are pairwise disjoint for each fixed  $j$ , and the same is true for the  $C_{j,s}$ ,  $A'_{j,s}$ , and  $C'_{j,s}$ . Thus, positive measure may occur only for a countable subset of values of  $s$ . Remark 6.7 ensures that the latter conclusion remains valid after the trimming. In conclusion, every  $(z, z') \in Y \times Y'$  is a continuity point of  $(y, y') \mapsto \tilde{\eta}_{y,y',s}$  for all but countably many values of  $s \in [0, 1]$ . Then, using Lemma 6.8,

$$\tilde{\eta}_{y,y'} = \int_0^1 \tilde{\eta}_{y,y',s} ds$$

is a coupling of  $\eta_y$  and  $\eta'_{y'}$ , depending continuously on  $y$  and  $y'$  and vanishing on a uniform neighborhood of  $K$ .  $\square$

**Remark 6.10.** When  $X = X'$  and the set  $K \subset X \times X$  is symmetric, one may exchange  $A_j, C_j$  with  $A'_j, C'_j$  in (6.9), and  $A_{j,s}, C_{j,s}$  with  $A'_{j,s}, C'_{j,s}$  in (6.11). Thus (recall Remark 6.5) when the roles of  $\eta$  and  $\eta'$  are exchanged the coupling  $\tilde{\eta}$  is replaced with  $\iota_* \tilde{\eta}$  in Proposition 6.6 and each  $\tilde{\eta}_{y',y}$  is replaced with  $\iota_* \tilde{\eta}_{y,y'}$  in Proposition 6.9. In particular, if  $\eta = \eta'$  and  $Y = Y'$  then the self-coupling  $\tilde{\eta}$  in Proposition 6.6 is symmetric, and the family  $\tilde{\eta}$  in Proposition 6.9 is involution-invariant, meaning that  $\tilde{\eta}_{y',y} = \iota_* \tilde{\eta}_{y,y'}$  for every  $(y, y') \in Y \times Y$ .

The following special case of Proposition 6.6 will be useful:

**Corollary 6.11.** *Let  $X = X'$  and assume that  $\eta$  and  $\eta'$  satisfy  $\eta(\{x\}) < \eta(X \setminus \{x\})$  and  $\eta'(\{x\}) < \eta'(X \setminus \{x\})$  for every  $x \in X$ . Then there exists a coupling  $\tilde{\eta}$  of  $\eta$  and  $\eta'$  that vanishes on a neighborhood of the diagonal of  $X \times X$ . If  $\eta = \eta'$  then the self-coupling  $\tilde{\eta}$  may be chosen to be symmetric.*

**6.2. Markov operators.** Let  $X$  be a metric space. We denote by  $\mathcal{B}(X)$  the Banach space of bounded measurable functions on  $X$ , with the norm

$$\|\psi\| = \sup\{|\psi(x)| : x \in X\}.$$

A *Markov operator* is a linear operator  $\mathcal{T} : \mathcal{B}(X) \rightarrow \mathcal{B}(X)$  of the form

$$(6.12) \quad \mathcal{T}\psi(x) = \int_X \psi(y) d\sigma_x(y)$$

where  $\{\sigma_x : x \in X\}$  is a measurable family of probability measures on  $X$ . It is clear that  $\mathcal{T}$  is a bounded operator, with  $\|\mathcal{T}\| = 1$ . We call it *continuous* if the map  $x \mapsto \sigma_x$  is continuous relative to the weak\* topology. Then  $\mathcal{T}$  preserves the subspace  $\mathcal{C}(X)$  of bounded continuous functions.

The *dual operator*  $\mathcal{T}^*$  is defined on the space of bounded finitely additive signed measures  $\eta$  on  $X$  with the total variation norm (see [29, IV.4.5]) by

$$(6.13) \quad \mathcal{T}^*\eta = \int_X \sigma_x d\eta(x).$$

The two are related by

$$(6.14) \quad \int_X \psi d(\mathcal{T}^*\eta) = \int_X (\mathcal{T}\psi) d\eta \text{ for every } \psi \text{ and } \eta.$$

A measure  $\eta$  is  $\mathcal{T}$ -invariant if  $\mathcal{T}^*\eta = \eta$ . If  $\mathcal{T}$  is continuous then  $\mathcal{T}^*$  preserves the subspace of bounded (countably additive) signed measures on  $X$ .

**Remark 6.12.** If  $\{\sigma_x : x \in X\}$  is such that every  $\sigma_x$  is a generic measure then, cf. Remark 5.1, the range of the associated dual Markov operator  $\mathcal{T}^*\eta = \int_X \sigma_x d\eta(x)$  is contained in the space of generic measures. In particular, any  $\mathcal{T}$ -invariant measure is a generic measure.

Suppose that  $X$  comes with a transitive  $G$ -action  $(g, x) \mapsto gx$ . Grassmannian manifolds  $\text{Gr}(r, d)$  and flag varieties  $\mathcal{F}(r, d)$  are the examples we have in mind. Then, to any probability measure  $\nu$  on  $G$  we may associate the Markov operators  $\mathcal{P}_\nu$  and  $\mathcal{P}_\nu^*$  in (3.1), (3.2), and (5.1)

$$(6.15) \quad \mathcal{P}_\nu\psi(x) = \int_G \psi(gx) d\nu(g) \text{ and } \mathcal{P}_\nu^*\eta = \int_G (g_*\eta) d\nu(g).$$

This corresponds to (6.12) with  $\sigma_x =$  the push-forward of  $\nu$  under the map  $g \mapsto gx$ . Note that  $\mathcal{P}_\nu$  is continuous if the  $G$ -action is continuous, and a measure  $\eta$  on  $Z$  is  $\mathcal{P}_\nu$ -invariant precisely if it is  $\nu$ -stationary. These are the fundamental examples of Markov operators in our context, but we will have to deal with other types as well.

One reason is that the kind of conclusion we are seeking, namely, that stationary measures give small weights to a neighborhood of the equator is local in nature. That is consistent with the fact that the information on the dynamics we can extract from Proposition 5.5 is clearly local. In contrast, the property of being a stationary measure is *not* local: the restriction of a  $\mathcal{P}_\nu$ -invariant measure  $\eta$  to some set  $U \subset X$  is usually not a  $\mathcal{P}_\nu$ -invariant measure. The way we handle this is by finding a “localized” Markov operator, related to the original one and to the domain  $U$  in an explicit manner, with respect to which the restriction  $\eta|_U$  is indeed an invariant measure.

**Remark 6.13.** The assumption that the  $G$ -action on  $X$  is transitive means that  $G \rightarrow X$ ,  $g \mapsto gx$  is surjective for any fixed  $x \in X$ . Then every probability measure  $\sigma$  on  $X$  lifts (non-uniquely) to a probability measure  $\mu$  on  $G$ : use the Hahn–Banach and Riesz–Markov theorems. In particular, (6.12) may be written in the form

$$\mathcal{T}\psi(x) = \int_G \psi(gx) d\mu_x(g)$$

for some family  $\{\mu_x : x \in X\}$  of probability measures on  $G$ . This general statement is not used in the present paper, but an explicit construction in a special case will appear in Lemma 16.1.

**Proposition 6.14.** *Let  $\mathcal{T} : \mathcal{B}(X) \rightarrow \mathcal{B}(X)$  be a Markov operator,  $\eta$  be a  $\mathcal{T}$ -invariant measure, and  $U \subset X$  be such that  $\eta(U) > 0$ . Then there exists a Markov operator  $\mathcal{T}_U : \mathcal{B}(U) \rightarrow \mathcal{B}(U)$  that leaves invariant the normalized restriction  $\eta_U$  of  $\eta$  to the subset  $U$ .*

*Proof.* We are going to find  $\{\sigma_{U,x} : x \in U\}$  such that  $\mathcal{T}_U \psi(x) = \int_U \psi(y) d\sigma_{U,x}(y)$  preserves  $\eta_U$ . Let  $\chi_U$  denote the characteristic function of  $U$ . Since  $\eta$  is  $\mathcal{T}$ -invariant,

$$\begin{aligned} 0 &= \int_X (\mathcal{T}\chi_U - \chi_U) d\eta = \int_X \left[ \int_X \chi_U(y) d\sigma_x(y) - \chi_U(x) \right] d\eta(x) \\ &= \int_U [\sigma_x(U) - 1] d\eta(x) + \int_{U^c} \sigma_x(U) d\eta(x), \end{aligned}$$

that is,

$$(6.16) \quad \int_U \sigma_x(U^c) d\eta(x) = \int_{U^c} \sigma_x(U) d\eta(x).$$

Let  $J$  be this number. If  $J = 0$ , there is not much to do:  $\eta_U$  turns out to be  $\mathcal{T}$ -invariant, and it suffices to take  $\sigma_{U,x} = \sigma_x$  for  $x \in U$ . If  $J > 0$ , define

$$(6.17) \quad \sigma_{U,x} = (\sigma_x | U) + \sigma_x(U^c) \frac{1}{J} \int_{U^c} (\sigma_z | U) d\eta(z)$$

for each  $x \in U$ . In other words,

$$\begin{aligned} \mathcal{T}_U \psi(x) &= \int_U \psi(y) d\sigma_x(y) + \sigma_x(U^c) \frac{1}{J} \int_{U^c} \int_U \psi(y) d\sigma_z(y) d\eta(z) \\ \text{and } \mathcal{T}_U^* \xi &= \int_U (\sigma_x | U) d\xi(x) + \int_U \sigma_x(U^c) d\xi(x) \frac{1}{J} \int_{U^c} (\sigma_z | U) d\eta(z) \end{aligned}$$

Observe that  $\sigma_{U,x}$  is a probability on  $U$ :

$$\sigma_{U,x}(U) = \sigma_x(U) + \sigma_x(U^c) \frac{1}{J} \int_{U^c} \sigma_x(U) d\xi(x) = \sigma_x(U) + \sigma_x(U^c) = 1.$$

Moreover, by the definition of  $J$ ,

$$\begin{aligned} \mathcal{T}^* \eta_U &= \int_U (\sigma_x | U) d\eta_U(x) + \int_U \sigma_x(U^c) d\eta_U(x) \frac{1}{J} \int_{U^c} (\sigma_z | U) d\eta(z) \\ &= \frac{1}{\eta(U)} \left( \int_U (\sigma_x | U) d\eta(x) + \int_{U^c} (\sigma_z | U) d\eta(z) \right) = \frac{1}{\eta(U)} \int_X (\sigma_x | U) d\eta(x). \end{aligned}$$

Each  $\psi \in \mathcal{B}(U)$  may be viewed as an element of  $\mathcal{B}(X)$  that vanishes outside  $U$ . Then

$$\begin{aligned} \int_U \psi d(\mathcal{T}_U^* \eta_U) &= \frac{1}{\eta(U)} \int_X \int_U \psi(y) d\sigma_x(y) d\eta(x) = \frac{1}{\eta(U)} \int_X \int_X \psi(y) d\sigma_x(y) d\eta(x) \\ &= \frac{1}{\eta(U)} \int_X \mathcal{T}\psi(x) d\eta(x) = \frac{1}{\eta(U)} \int_X \psi(x) d\eta(x) = \int_U \psi d\eta_U. \end{aligned}$$

Thus,  $\mathcal{T}^* \eta_U = \eta_U$ , as we wanted to prove.  $\square$

The operator  $\mathcal{T}_U$  in Proposition 6.14 need not be continuous, in general. In the next proposition we fix that problem.

**Proposition 6.15.** *Let  $\mathcal{T} : \mathcal{B}(X) \rightarrow \mathcal{B}(X)$  be a Markov operator,  $\eta$  be a  $\mathcal{T}$ -invariant measure, and  $U \subset X$  be such that  $\eta(U) > 0$ . Assume that  $x \mapsto \sigma_x$  is continuous on  $U$ . Then there is a continuous Markov operator  $\mathcal{T}_U : \mathcal{B}(U) \rightarrow \mathcal{B}(U)$  such that the normalized restriction  $\eta_U$  is  $\mathcal{T}_U$ -invariant.*



*Proof.* The strategy is to consider a monotone family  $\{U_t : t \in [0, 1]\}$  of subsets of  $U$  with pairwise disjoint boundaries, and to associate to each  $U_t$  a Markov operator  $\mathcal{T}_t$  such that  $\mathcal{T}_t^* \eta_U = \eta_U$ . These operators will still be discontinuous, but we can get rid of the discontinuities by integrating with respect to  $t$ . The details follow.

As before, let  $J$  be the number in (6.16). When  $J = 0$  there is nothing to do, because in that case the construction in Proposition 6.14 does yield a continuous Markov operator  $\mathcal{T}_U$ . From now on, assume that  $J > 0$ . Let  $a > 0$  be a small number. For each  $t \in [0, 1]$ , define  $U_t = \{x \in U : d(x, U^c) \geq at\}$  and

$$J(t) = \int_U \sigma_x(U_t^c) d\eta(x).$$

Note that  $J(t) \geq J > 0$ . Then let  $\xi_t$  be the probability measure defined on  $U$  by

$$(6.18) \quad \int_U (\sigma_x | U_t) d\eta(x) + J(t)\xi_t = \eta | U.$$

Observe that  $\xi_t$  is well defined (each  $\sigma_x | U_t$  may be viewed as a measure on  $U$ , since  $U_t \subset U$ , and so all the terms in this identity are measures on  $U$ ) and it is indeed a probability measure.

Now let  $\mathcal{T}_t : \mathcal{B}(U) \rightarrow \mathcal{B}(U)$  be the Markov operator associated to the family

$$(6.19) \quad \sigma_{x,t} = (\sigma_x | U_t) + \sigma_x(U_t^c)\xi_t$$

of probability measures on  $U$ . Condition (6.18) means that  $\eta_U$  is  $\mathcal{T}_t$ -invariant:

$$(6.20) \quad \begin{aligned} \mathcal{T}_t^* \eta_U &= \int_U (\sigma_x | U_t) d\eta_U(x) + \int_U \sigma_x(U_t^c)\xi_t d\eta_U(x) \\ &= \int_U (\sigma_x | U_t) d\eta_U(x) + \frac{J(t)}{\eta(U)}\xi_t = \eta_U. \end{aligned}$$

Next, define  $\mathcal{T}_U$  to be the Markov operator associated to the family of probability measures

$$(6.21) \quad \sigma_{U,x} = \int_0^1 \sigma_{x,t} dt.$$

It is clear from (6.20) that  $\eta_U$  is  $\mathcal{T}_U$ -invariant.

We are left to show that the map  $x \mapsto \sigma_{U,x}$  is continuous with respect to the weak\* topology, that is, that

$$(6.22) \quad x \mapsto \int \varphi d\sigma_{U,x} = \int_0^1 \left( \int_{U_t} \varphi(y) d\sigma_x(y) + \sigma_x(U_t^c) \int_U \varphi(z) d\xi_t(z) \right) dt$$

is continuous for any bounded continuous function  $\varphi : U \rightarrow \mathbb{R}$ . This will be a consequence of the following fact:

**Lemma 6.16.** *Let  $Z$  be a metric space,  $\sigma_0$  be a probability measure on  $Z$  and  $g : Z \rightarrow \mathbb{R}$  be a measurable function such that the closure of the set of discontinuity points has zero  $\sigma_0$ -measure. Then  $\sigma_0$  is a continuity point of the map  $\sigma \mapsto \int_Z g d\sigma$  in the space of probability measures on  $Z$  with the weak\* topology.*

*Proof.* Denote by  $R$  the closure of the set of discontinuity points. Given  $\varepsilon > 0$ , let  $V$  be an open neighborhood of  $R$  whose closure  $\bar{V}$  satisfies  $\sigma_0(\bar{V}) < \varepsilon$ . Then  $\sigma(\bar{V}) < \varepsilon$  for any  $\sigma$  in a weak\* neighborhood of  $\sigma_0$ . By the Tietze extension theorem, there

exists a continuous function  $h : Z \rightarrow \mathbb{R}$  coinciding with  $g$  on the complement of  $V$  and satisfying  $\|h\| \leq \|g\|$ . Then

$$\left| \int_Z g d\sigma - \int_Z g d\sigma_0 \right| \leq \left| \int_Z h d\sigma - \int_Z h d\sigma_0 \right| + 4\|g\|\varepsilon \leq (1 + 4\|g\|)\varepsilon$$

for any  $\sigma$  in a weak\* neighborhood of  $\sigma_0$ . Thus,  $\sigma_0$  is a continuity point.  $\square$

Going back to proving the proposition, fix any  $z \in U$  and consider  $Z = X$  and  $\sigma_0 = \sigma_z$ . Keep in mind that  $x \mapsto \sigma_x$  is assumed to be continuous on  $U$ . Thus, applying Lemma 6.16 twice, to  $g = \varphi\chi_{U_t}$  and to  $g = \chi_{U_t^c}$ , we see that  $z$  is a point of discontinuity of

$$(6.23) \quad x \mapsto \int_{U_t} \varphi(y) d\sigma_x(y) + \sigma_x(U_t^c) \int_U \varphi(z) d\xi_t(z)$$

only if the boundary  $\partial U_t = \partial U_t^c$  has positive measure for  $\sigma_x$ . Since these boundaries are pairwise disjoint when  $t$  varies, the latter can only happen for countably many values of  $t$ . Thus, we may apply Lemma 6.8 to  $Z = X$  and the function  $g(x, t)$  given by the right-hand side of (6.23), to conclude that (6.22) is continuous.  $\square$

**Remark 6.17.** The localization procedure in Propositions 6.14 and 6.15 preserves the class of non-atomic measures and, when  $X$  is an algebraic variety, also the class of generic measures. That is a direct consequence of (6.17), (6.18), (6.19), and Remark 5.1.

**6.3. Invariant couplings.** Let  $X$  be a metric space  $X$  and  $\mathcal{T} : \mathcal{B}(X) \rightarrow \mathcal{B}(X)$  be a Markov operator, given by

$$\mathcal{T}\psi(x) = \int_X \psi(y) d\sigma_x(y).$$

A *self-coupling* of  $\mathcal{T}$  is a Markov operator  $\tilde{\mathcal{T}} : \mathcal{B}(X \times X) \rightarrow \mathcal{B}(X \times X)$  of the form

$$(6.24) \quad \tilde{\mathcal{T}}\tilde{\psi}(x, x') = \int_{X \times X'} \tilde{\psi}(y, y') d\tilde{\sigma}_{x, x'}(y, y')$$

where each  $\tilde{\sigma}_{x, x'}$  is a coupling of  $\sigma_x$  and  $\sigma_{x'}$ . The self-coupling  $\tilde{\mathcal{T}}$  is *continuous* if the map  $(x, x') \mapsto \tilde{\sigma}_{x, x'}$  is continuous on  $X \times X$ .

**Lemma 6.18.** *If  $\tilde{\eta}$  is a coupling of  $\eta$  and  $\eta'$  and  $\tilde{\mathcal{T}}$  is a self-coupling of  $\mathcal{T}$  then  $\tilde{\mathcal{T}}^*\tilde{\eta}$  is a coupling of  $\mathcal{T}^*\eta$  and  $\mathcal{T}^*\eta'$ .*

*Proof.* Let  $\tilde{\psi} : X \times X \rightarrow \mathbb{R}$  be any bounded measurable function that depends only on the first variable:  $\tilde{\psi}(x, x') = \psi(x)$  for some  $\psi \in \mathcal{B}(X)$ . By definition,

$$\begin{aligned} \int_{X \times X} \tilde{\psi} d(\tilde{\mathcal{T}}^*\tilde{\eta}) &= \int_{X \times X} \int_{X \times X} \tilde{\psi}(y, y') d\tilde{\sigma}_{x, x'}(y, y') d\tilde{\eta}(x, x') \\ &= \int_{X \times X} \int_{X \times X} \psi(y) d\tilde{\sigma}_{x, x'}(y, y') d\tilde{\eta}(x, x'). \end{aligned}$$

Since  $\tilde{\sigma}_{x, x'}$  projects to  $\sigma_x$  and  $\tilde{\eta}$  projects to  $\eta$  on the first factor, this last expression may be written as

$$\int_{X \times X} \int_X \psi(y) d\sigma_x(y) d\tilde{\eta}(x, x') = \int_X \int_X \psi(y) d\sigma_x(y) d\eta(x) = \int_X \psi d(\mathcal{T}^*\eta).$$

This proves that  $\tilde{\mathcal{T}}^*\tilde{\eta}$  projects to  $\mathcal{T}^*\eta$  on the first factor. Analogously, it projects to  $\mathcal{T}^*\eta'$  on the second factor.  $\square$

**Lemma 6.19.** *Assume that  $X$  is compact and  $\tilde{\mathcal{T}}$  is continuous, and let  $\eta$  and  $\eta'$  be  $\mathcal{T}$ -invariant probability measures on  $X$ . Given any coupling  $\tilde{\eta}_0$  of  $\eta$  and  $\eta'$ , every accumulation point  $\tilde{\eta}$  of the sequence*

$$\tilde{\eta}_n = \frac{1}{n} \sum_{j=0}^{n-1} \tilde{\mathcal{T}}^{n*} \tilde{\eta}_0$$

is a  $\tilde{\mathcal{T}}$ -invariant coupling of  $\eta$  and  $\eta'$ . In particular,  $\tilde{\mathcal{T}}$ -invariant couplings do exist.

*Proof.* By Lemma 6.18, every  $\tilde{\eta}_n$  is a coupling of  $\eta$  and  $\eta'$ . By compactness, there exists  $(n_i)_i \rightarrow \infty$  such that  $(\tilde{\eta}_{n_i})_i$  converges to some  $\tilde{\eta}$  in the weak\* topology. Clearly,  $\tilde{\eta}$  is still a coupling of  $\eta$  and  $\eta'$ . Let  $\tilde{\psi} : X \times X \rightarrow \mathbb{R}$  be any bounded continuous function. The assumption ensures that  $\tilde{\mathcal{T}}\tilde{\psi}$  is also continuous. Thus,

$$\begin{aligned} \int_{X \times X} (\tilde{\mathcal{T}}\tilde{\psi} - \tilde{\psi}) d\tilde{\eta} &= \lim_i \int_{X \times X} (\tilde{\mathcal{T}}\tilde{\psi} - \tilde{\psi}) d\tilde{\eta}_{n_i} \\ &= \lim_i \int_{X \times X} \frac{1}{n_i} (\tilde{\mathcal{T}}^{n_i}\tilde{\psi} - \tilde{\psi}) d\tilde{\eta}_0 = 0 \end{aligned}$$

(recall that  $\|\tilde{\mathcal{T}}^n\tilde{\psi}\| \leq \|\tilde{\psi}\|$  for every  $n$ ). Thus,  $\tilde{\mathcal{T}}^*\tilde{\eta} = \tilde{\eta}$ , as we wanted to prove.  $\square$

**Remark 6.20.** A self-coupling  $\tilde{\mathcal{T}}$  is *symmetric* if  $\sigma_{x',x} = \iota_*\sigma_{x,x'}$  for all  $x, x' \in X$ . If  $\tilde{\mathcal{T}}$  is symmetric and  $\tilde{\eta}$  is a symmetric self-coupling of  $\eta'$  then  $\tilde{\mathcal{T}}^*\tilde{\eta}$  is a symmetric self-coupling of  $\tilde{\mathcal{T}}^*\eta$ . Moreover, the  $\tilde{\mathcal{T}}$ -invariant self-coupling  $\tilde{\eta}$  in Lemma 6.19 may be taken to be symmetric.

**6.4. Margulis functions.** As before, let

$$\mathcal{T} : \mathcal{B}(X) \rightarrow \mathcal{B}(X), \quad \mathcal{T}\psi(x) = \int_X \psi(y) d\sigma_x(y)$$

be a Markov operator on a metric space  $X$ . Let  $X = A \cup B$  be a partition of  $X$  into disjoint sets  $A$  and  $B$ . An (*additive*) *Margulis function* for  $\mathcal{T}$  relative to  $(A, B)$  is a measurable function  $\Psi : X \rightarrow [0, \infty]$  such that there exist  $\kappa_A > 0$  and  $\kappa_B > 0$  such that

$$(6.25) \quad \mathcal{T}\Psi(x) \leq \Psi(x) - \kappa_A \text{ for every } x \in A$$

$$(6.26) \quad \mathcal{T}\Psi(x) \leq \Psi(x) + \kappa_B \text{ for every } x \in B.$$

( $\Psi$  is usually not bounded, but its image under  $\mathcal{T}$  is easily defined using monotone convergence: let  $\mathcal{T}\Psi = \lim_n \mathcal{T}(\min\{\Psi, n\})$ . Then  $\int_X \mathcal{T}\Psi d\zeta = \int \Psi d(\mathcal{T}^*\zeta)$  for any probability measure  $\zeta$  on  $X$ .) We make following technical assumption, which is used in the context of (6.30): there exists  $L > 0$  such that  $\Psi$  is lower semi-continuous on  $\Psi^{-1}([L, \infty])$ .

Margulis functions are a very effective tool for estimating the spatial distribution of  $\mathcal{T}$ -invariant measures. The simple lemma that follows illustrates this idea:

**Lemma 6.21.** *Let  $\Psi : X \rightarrow [0, \infty]$  be a Margulis function for a Markov operator  $\mathcal{T}$  relative to a partition  $(A, B)$ . Let  $\zeta$  be any measure on  $X$  such that  $\int \Psi d\zeta < \infty$  and  $\int_X \mathcal{T}\Psi d\zeta \geq \int_X \Psi d\zeta$ . Then*

$$(6.27) \quad \zeta(B) \geq \frac{\kappa_A}{\kappa_A + \kappa_B} \zeta(X).$$

*In particular, this holds if  $\zeta$  is  $\mathcal{T}$ -invariant and satisfies  $\int \Psi d\zeta < \infty$ .*

*Proof.* We have

$$\int_X \Psi d\zeta \leq \int_X \mathcal{T}\Psi d\zeta \leq \int_X \Psi d\zeta - \kappa_A \zeta(A) + \kappa_B \zeta(B).$$

Thus,  $\kappa_A \zeta(A) - \kappa_B \zeta(B) \leq 0$ , which is the same as (6.27). To get the last claim, just note that  $\int_X \mathcal{T}\Psi d\zeta = \int_X \Psi d\zeta$  if  $\zeta$  is  $\mathcal{T}$ -invariant.  $\square$

**Remark 6.22.** Given a set  $Y \subset X$ , we call *multiplicative Margulis function* for  $\mathcal{T}$  relative to  $(X, Y)$  any measurable function  $\Phi : X \rightarrow [1, \infty]$  such that  $\log \Phi$  is uniformly continuous,  $\Phi(x) = \infty$  if and only if  $x \in Y$ ,  $\Phi$  is a proper function on  $X \setminus Y$ , and there exist constants  $c < 1$  and  $b < \infty$  such that

$$(6.28) \quad \mathcal{T}\Phi(x) \leq c\Phi(x) + b \text{ for all } x \in X.$$

If  $\Phi$  is a multiplicative Margulis function then  $\log \Phi$  is an additive Margulis function relative to the partition  $(A, B)$  defined by

$$A = \{x \in X : \Phi(x) > \alpha\} \text{ and } B = \{x \in X : \Phi(x) \leq \alpha\}$$

for any  $\alpha > b/(1-c)$ . Indeed, the Jensen inequality implies that

$$\mathcal{T} \log \Phi(x) \leq \log \mathcal{T}\Phi(x) \leq \log(c\Phi(x) + b) \text{ for every } x.$$

Moreover,

$$\log(cy + b) \leq \begin{cases} \log y + \log(c + b) & \text{for every } y \geq 1 \\ \log y + \log(c + b/\alpha) & \text{if } y > \alpha. \end{cases}$$

Thus, we may take  $\kappa_A = -\log(c + b/\alpha) > 0$  and any  $\kappa_B \geq \log(c + b)$ . On the other hand, as was already pointed out in the Introduction, it is not true that if  $\Psi$  is an additive Margulis function then  $\exp \Psi$  is a multiplicative one.

**Proposition 6.23.** *Assume that  $X$  is compact and let  $\tilde{\mathcal{T}}$  be a continuous self-coupling of  $\mathcal{T}$ . Let  $\Psi : X \times X \rightarrow [0, +\infty]$  be a Margulis function for  $\tilde{\mathcal{T}}$  which is bounded outside every neighborhood of the diagonal, and let  $\eta$  be a non-atomic  $\mathcal{T}$ -invariant measure on  $X$ . Then one can find a  $\tilde{\mathcal{T}}$ -invariant self-coupling  $\tilde{\eta}$  of  $\eta$  and a sequence  $(\tilde{\eta}_j)_j$  of self-couplings of  $\eta$  converging to  $\tilde{\eta}$  in the weak\* topology and such that*

$$(6.29) \quad \int_{X \times X} \Psi d\tilde{\eta}_j < \infty \text{ and } \int_{X \times X} \tilde{\mathcal{T}}\Psi d\tilde{\eta}_j \geq \int_{X \times X} \Psi d\tilde{\eta}_j \text{ for every } j.$$

*Proof.* By Corollary 6.11, there exists some self-coupling  $\hat{\eta}_0$  of  $\eta$  that vanishes on a neighborhood of the diagonal. Then  $\int_{X \times X} \Psi d\hat{\eta}_0$  is finite. Conditions (6.25) and (6.26) imply that  $\tilde{\mathcal{T}}^j \Psi(x) \leq \Psi(x) + j\kappa_B$  for every  $x \in X$ , and so

$$\int_{X \times X} \tilde{\mathcal{T}}^j \Psi d\hat{\eta}_0 \leq \int_{X \times X} \Psi d\hat{\eta}_0 + j\kappa_B < \infty$$

for every  $j$ . Let  $\tilde{\eta} = \lim_i \hat{\eta}_{n_i}$  be any weak\* accumulation point of the sequence

$$\hat{\eta}_n = \frac{1}{n} \sum_{j=0}^{n-1} \tilde{\mathcal{T}}^{n*} \hat{\eta}_0.$$

As noted in Lemma 6.19, every  $\hat{\eta}_n$  is a self-coupling of  $\eta$  with  $\int_{X \times X} \Psi d\hat{\eta}_n < \infty$ , and  $\tilde{\eta}$  is a  $\tilde{\mathcal{T}}$ -invariant self-coupling of  $\eta$ .

If  $\int_X \Psi d\tilde{\eta}$  is finite then the claim follows by taking  $\tilde{\eta}_j = \tilde{\eta}$  for every  $j$ . In this case the equality holds on the second part of (6.29). Now suppose that  $\int_X \Psi d\tilde{\eta}$  is infinite. By the lower semi-continuity assumption on  $\Psi$ , this implies that

$$(6.30) \quad \int_X \Psi d\hat{\eta}_{m_i} \rightarrow \infty \text{ as } i \rightarrow \infty.$$

Then  $\int_X \Psi d\left(\tilde{\mathcal{T}}^{*n}\hat{\eta}_0\right)$  must be unbounded. In particular, one can find  $(m_j)_j \rightarrow \infty$  such that

$$\int_{X \times X} \tilde{\mathcal{T}}\Psi d\hat{\eta}_{m_j} - \int_{X \times X} \Psi d\hat{\eta}_{m_j} = \frac{1}{m_j} \left( \int_{X \times X} \Psi d\left(\tilde{\mathcal{T}}^{*m_j}\hat{\eta}_0\right) - \int_{X \times X} \Psi d\hat{\eta}_0 \right) \geq 0.$$

Thus, it suffices to take  $\tilde{\eta}_j = \hat{\eta}_{m_j}$  for every  $j$ .  $\square$

**6.5. Adapted operators.** Let  $Z$  be an algebraic variety endowed with a continuous  $G$ -action  $(g, z) \mapsto gz$ . Let  $\nu$  be a compactly supported probability measure on  $G$ , and  $\eta$  be a  $\nu$ -stationary probability measure on  $Z$ . For each  $z \in Z$ , let  $\nu_z$  denote the push-forward of  $\nu$  under the map  $g \mapsto gz$ .

Let  $X \subset Z$  and  $a_X > 0$  be some small number. Consider the family of subsets  $X_t = \{x \in X : d(x, Z \setminus X) \geq a_X t\}$ ,  $t \in [0, 2]$ . Note that  $t \mapsto X_t$  is monotone decreasing. Define the  $\nu$ -core of  $X$  to be

$$(6.31) \quad \mathcal{X}_\nu X = \{x \in X_2 : gx \in X_2 \text{ for all } g \in \text{supp } \nu\}.$$

The complement  $\partial_\nu X = X \setminus \mathcal{X}_\nu X$  is called the  $\nu$ -border of  $X$ . Define also

$$(6.32) \quad \mathcal{X}_\nu^\# X = \{x \in \mathcal{X}_\nu X : g^{-1}x \in \mathcal{X}_\nu X \text{ for all } g \in \text{supp } \nu\}.$$

It is clear that  $\mathcal{X}_\nu^\# X \subset \mathcal{X}_\nu X \subset X_2 \subset X$ . Moreover,

$$(6.33) \quad \nu_x(X \setminus X_2) = 0 \text{ for all } x \in \mathcal{X}_\nu X, \text{ and}$$

$$(6.34) \quad \nu_x(\mathcal{X}_\nu^\# X) = 0 \text{ for all } x \in \partial_\nu X.$$

**Example 6.24.** In our initial applications,  $Z$  will be a Grassmannian manifold  $\text{Gr}(r, d)$  and  $X$  will be the closed  $\varepsilon$ -neighborhood  $E_r(\varepsilon)$  of the subset of  $r$ -dimensional subspaces of  $\mathbb{R}^d$  contained in the equator  $E$ . Later we will also take  $Z$  to be a flag variety  $\mathcal{F}(r, d)$  and  $X$  to be the closed subset  $E_r^\circ(\varepsilon)$  of flags whose  $r$ -coordinate  $F_r$  is in  $E_r(\varepsilon)$ . We will always take  $a_X = \varepsilon/100$ , which means that  $X_t = E_r((1-t/100)\varepsilon)$  and  $X_t = E_r^\circ((1-t/100)\varepsilon)$ , respectively, for all  $t \in [0, 2]$ .

A Markov operator  $\mathcal{T} : \mathcal{B}(X) \rightarrow \mathcal{B}(X)$  is said to be *adapted to*  $(\nu, X)$  if the associated family of probability measures  $\{\sigma_x : x \in X\}$  satisfies

- (a)  $\sigma_x = \nu_x$  for every  $x$  in a neighborhood of the  $\nu$ -core  $\mathcal{X}_\nu X$ ;
- (b)  $\sigma_x(\mathcal{X}_\nu^\# X) = 0$  for every  $x$  in the  $\nu$ -border  $\partial_\nu X$ ;
- (c)  $\sigma_x$  is generic for every  $x \in X$ .

The assumption of the next proposition implies that  $x \mapsto \sigma_x$  is continuous on  $U$ , and so Proposition 6.15 does hold in this setting. The assumption is satisfied, in particular, if the Markov operator  $\mathcal{T}$  is adapted to  $(\nu, X)$  and  $U$  is a subset of the  $\nu$ -core of  $X$ .

**Proposition 6.25.** *Let  $\mathcal{T} : \mathcal{B}(X) \rightarrow \mathcal{B}(X)$  be a Markov operator,  $\eta$  be a  $\mathcal{T}$ -invariant measure, and  $U \subset X$  be such that  $\eta(U) > 0$ . Suppose that*

- (a)  $\sigma_x = \nu_x$  for every  $x \in U$ ;
- (b)  $\sigma_x(\mathcal{X}_\nu^\# U) = 0$  for every  $x \notin U$ ;

(c)  $\sigma_x$  is a generic measure for every  $x \in X$ .

Then the continuous Markov operator  $\mathcal{T}_U : \mathcal{B}(U) \rightarrow \mathcal{B}(U)$  given by Proposition 6.15 is adapted to  $(\nu, U)$ .

*Proof.* Let  $U^c = X \setminus U$  and  $U_t^c = X \setminus U_t$ . Recall that  $\mathcal{T}_U$  is given by the probability measures  $\sigma_{U,x}$  defined in (6.21). Since  $\eta = \int_X \sigma_y d\eta(y)$ , because  $\eta$  is assumed to be  $\mathcal{T}$ -invariant, the definition (6.18) of  $\xi_t$  means that

$$\xi_t = \frac{1}{J(t)} \left( \int_{U^c} (\sigma_y | U) d\eta(y) + \int_U (\sigma_y | U \cap U_t^c) d\eta(y) \right),$$

and so

$$(6.35) \quad \sigma_{x,t} = (\sigma_x | U_t) + \frac{\sigma_x(U_t^c)}{J(t)} \left( \int_{U^c} (\sigma_y | U) d\eta(y) + \int_U (\sigma_y | U \cap U_t^c) d\eta(y) \right).$$

Condition (c) in the definition of an adapted operator follows directly from Remark 6.17. Let us check conditions (a) and (b).

If  $x \in \mathcal{X}_\nu U$  then  $gx \in U_2$  for every  $g \in \text{supp } \nu$ . Since  $U_1$  is a neighborhood of  $U_2$ , and the support of  $\nu$  is compact, it follows that there exists a neighborhood  $V$  of  $\mathcal{X}_\nu U$  such that  $gx \in U_1$  for every  $x \in V$  and  $g \in \text{supp } \nu$ . Thus  $\sigma_x(U_1^c) = \nu_x(U_1^c) = 0$ , and so, for all  $t \in [0, 1]$ , the second term on the right hand side of (6.35) vanishes, which means that  $\sigma_{x,t} = \sigma_x = \nu_x$ . Integrating with respect to  $t$ , we find that  $\sigma_{U,x} = \sigma_x = \nu_x$  for all  $x \in V$ , which proves condition (a).

Now consider  $x \in \partial_\nu U$ . We claim that all three terms on the right-hand side of (6.35) vanish on  $\mathcal{X}_\nu^\# U$ . Initially, (6.34) implies that  $\sigma_x(\mathcal{X}_\nu^\# U) = \nu_x(\mathcal{X}_\nu^\# U) = 0$ , which sets the claim for the first term. The case of the second term is an immediate consequence of the assumption that  $\sigma_y(\mathcal{X}_\nu^\# U) = 0$  for every  $y \in U^c$ . Finally,  $(\sigma_y | U \cap U_t^c)(\mathcal{X}_\nu^\# U) = 0$  for every  $y \in U$  and  $t \in [0, 1]$ , just because  $\mathcal{X}_\nu^\# U$  is contained in  $U_t$ . Hence, the third term is also zero on  $\mathcal{X}_\nu^\# U$ . This proves that  $\sigma_{x,t}(\mathcal{X}_\nu^\# U) = 0$  for all  $t \in [0, 1]$ , which implies that  $\sigma_{U,x}$  satisfies (b).  $\square$

**Remark 6.26.** By definition  $\nu_x(\mathcal{X}_\nu^\# U) = 0$  for every  $x \notin U$ . Compare conditions (a) and (b) in Proposition 6.25.

## 7. OUTLINE OF THE PROOF

Given any subspaces  $U$  and  $V$  of  $\mathbb{R}^d$ , define

$$(7.1) \quad d(U, V) = \sup_{u \in U} \inf_{v \in V} d(u, v) = \sup_{u \in U} \inf_{v \in V} |\sin \angle(u, v)|.$$

Then  $0 \leq d(U, V) \leq 1$ , with  $d(U, V) = 0$  if and only if  $U \subset V$  and  $d(U, V) = 1$  if and only if  $U \cap V^\perp \neq \{0\}$ . In particular,

$$(7.2) \quad \dim U > \dim V \quad \Rightarrow \quad d(U, V) = 1.$$

It is also clear that  $d(U_2, V_1) \geq d(U_1, V_2)$  whenever  $U_1 \subset U_2$  and  $V_1 \subset V_2$ .

The function  $d(\cdot, \cdot)$  in (7.1) is clearly not symmetric, in general. However, its restriction to each Grassmannian manifold  $\text{Gr}(r, d)$  turns out to be a distance, invariant under the action of the orthogonal group  $O(d)$  on the Grassmannian. The case  $r = 1$  is just (5.16).

Let  $E_r \subset \text{Gr}(r, d)$  be the set of  $r$ -dimensional subspaces contained in the equator  $E$ . In particular,  $\text{Gr}(1, d) = P$  and  $E_1 = E$ . For  $\varepsilon > 0$ , let  $E_r(\varepsilon) \subset \text{Gr}(r, d)$  be the closed  $\varepsilon$ -neighborhood of  $E_r$ . We also denote  $E_r(\varepsilon, \varepsilon') = E_r(\varepsilon) \setminus E_r(\varepsilon')$  for

$0 < \varepsilon' < \varepsilon$ . For each  $X = E_r(\varepsilon)$ , we always take  $a_X = \varepsilon/100$  in the definition (6.31) of the  $\nu$ -core  $\mathcal{X}_\nu X$  and the  $\nu$ -border  $\partial_\nu X$ .

**7.1. Main inductive statement.** We are going to prove Theorem B by contradiction: suppose that there exists a discontinuity point  $\nu_\infty$  for the largest Lyapunov exponent  $\lambda_1$ . Then, as we saw in Section 5 (Proposition 5.2 and Corollary 5.4), there exists a sequence  $(\nu_k)_k$  of generic measures with Zariski-dense supports converging to some  $\nu_\infty$  on  $G$  and such that their (unique) stationary measures  $(\eta_k)_k$  on  $P$  converge to a  $\nu_\infty$ -stationary measure  $\eta_\infty$  satisfying  $\eta_\infty(E) > 0$ .

**Theorem 7.1.** *Let  $1 \leq r \leq d$  and suppose that there exist constants  $\varepsilon_r > 0$  and  $n_r \in \mathbb{N}$  such that:*

(i) *For each  $k \in \mathbb{N}$  there is a continuous Markov operator*

$$\mathcal{T}_{k,r} : \mathcal{B}(E_r(\varepsilon_r)) \rightarrow \mathcal{B}(E_r(\varepsilon_r)), \quad \mathcal{T}_{k,r}\varphi(x) = \int_{E_r(\varepsilon_r)} \varphi(y) d\sigma_{k,r,x}(y)$$

*adapted to  $(\nu_k^{(n_r)}, E_r(\varepsilon_r))$ .*

(ii) *For each  $k \in \mathbb{N}$  there is a  $\mathcal{T}_{k,r}$ -invariant probability measure  $\eta_{k,r}$  on  $E_r(\varepsilon_r)$  such that  $\eta_{\infty,r} = \lim_k \eta_{k,r}$  exists and satisfies  $\eta_{\infty,r}(E_r) > 0$ .*

*If  $r < \dim E$  then there exist constants  $\varepsilon_{r+1}$  and  $K_{r+1} \in \mathbb{N}$  such that (i) and (ii) hold when  $r$  is replaced with  $r + 1$ . If  $r = \dim E$  then (i) and (ii) cannot happen.*

By Remark 6.12, the invariant measures  $\eta_{k,r}$  are automatically generic measures.

**Remark 7.2.** Since the support of  $\nu_k$  is Zariski-dense, it follows from Proposition 5.2 that for each  $1 \leq r \leq d$  there exists a unique  $\nu_k$ -stationary probability measure  $\hat{\eta}_{k,r}$  in  $\text{Gr}(r, d)$ . However, even if we assume that there exists a subsequence along which  $(\hat{\eta}_{k,1})_k$  converges to a measure  $\hat{\eta}_{\infty,1}$  with  $\hat{\eta}_{\infty,1}(E) > 0$ , it is not clear that for any  $r > 1$  the sequence  $(\hat{\eta}_{k,r})_k$  admits a subsequence converging to some measure  $\hat{\eta}_{\infty,r}$  with  $\hat{\eta}_{\infty,r}(E_r) > 0$ . Thus, Theorem 7.1 is not immediately obvious. Indeed, our measures  $\eta_{k,r}$  on  $E_r(\varepsilon_r)$  are *not* the normalized restrictions of the  $\hat{\eta}_{k,r}$ .

To deduce Theorem B, note that the assumptions of Theorem 7.1 hold for  $r = 1$ , with  $n_1 = 1$  and  $\varepsilon_1 = 1$ , so that  $E_1(\varepsilon_1) = P$ , and

$$\mathcal{T}_{k,1} = \mathcal{P}_{\nu_k} : \mathcal{B}(P) \rightarrow \mathcal{B}(P) \text{ and } \eta_{k,1} = \eta_k \text{ for } k \in \mathbb{N}.$$

Indeed, it is clear that  $\mathcal{P}_{\nu_k}$  is adapted to  $(\nu_k, P)$  and leaves  $\eta_k$  invariant. Recall that  $\nu_k$  is taken to be generic for  $k \in \mathbb{N}$ . Moreover, the limit  $\eta_{\infty,1} = \eta_\infty$  satisfies  $\eta_\infty(E) > 0$ . Then we can iterate Theorem 7.1 until we arrive at the case  $r = \dim E$ , which leads to a contradiction, which proves Theorem B.

The proof of Theorem 7.1 occupies what is left of this paper. In the remainder of the present section we outline the main ideas. Initially, we discuss the case  $r = 1$ , which involves many of the ingredients of the general step, though not all. Then we hint at how these ideas can be extended to  $r > 1$ .

Throughout,  $E_1(\varepsilon)^2 = E_1(\varepsilon) \times E_1(\varepsilon)$  and  $E_1(\varepsilon, \varepsilon')^2 = E_1(\varepsilon, \varepsilon') \times E_1(\varepsilon, \varepsilon')$  for any  $0 < \varepsilon' < \varepsilon$ . The following simple fact will be used a few times:

**Remark 7.3.** Given any  $\beta > 0$  and  $n \in \mathbb{N}$ , the neighborhood  $E_r(\alpha)$  is contained in the  $\nu_\infty^{(n)}$ -core of  $E_r(\beta/2)$  for all  $1 \leq r \leq d$  if  $\alpha$  is sufficiently small, depending only on  $\nu_\infty$ ,  $n$  and  $\beta$ . This is because the equator  $E$  is  $\nu_\infty$ -invariant. Since  $\text{supp } \nu_k^{(n)}$

converges to  $\text{supp } \nu_{\mathcal{O}}^{(n)}$ , it follows that if  $\alpha > 0$  is sufficiently small and  $k \in \mathbb{N}$  is large enough, depending only on  $\nu_{\mathcal{O}}$ ,  $n$  and  $\beta$ , then  $E_r(\alpha)$  is contained in the  $\nu_k^{(n)}$ -core of  $E_r(\beta)$  for all  $1 \leq r \leq d$ .

**7.2. First step.** Let us consider constants  $\varepsilon_1 > 0$  and  $n_1 \in \mathbb{N}$ , continuous Markov operators

$$\mathcal{T}_{k,1} : \mathcal{B}(E_1(\varepsilon_1)) \rightarrow \mathcal{B}(E_1(\varepsilon_1)), \quad \mathcal{T}_{k,1}\varphi(x) = \int_G \varphi(y) d\sigma_{k,1,x}(y)$$

and  $\mathcal{T}_{k,1}$ -invariant measures  $\eta_{k,1}$  as in Theorem 7.1. We may start from  $n_1 = 1$  and  $\varepsilon_1 = 1$ , as in the previous section, but along the way we replace the operators with convenient iterates, which means that  $n_1$  is increased, and we also localize them to suitable neighborhoods of the equator  $E$ , using Propositions 6.15 and 6.25 and Remark 7.3, which entails reducing  $\varepsilon_1$ .

The way we set this is by replacing  $n_1$  with a variable  $n \in \mathbb{N}$ , which we think of as a free parameter, and by taking  $\mathcal{T}_{k,1}$  and  $\varepsilon_1$  as functions of  $n$ . The conditions on  $\varepsilon_1$  are given in (7.3) and (9.12), depending on  $\nu_{\mathcal{O}}$ ,  $\delta$  and  $n$ . The condition on  $n$  is stated only near the end of the construction, in (11.13), depending on  $\nu_{\mathcal{O}}$  and  $\delta$ .

1. Up to reducing  $\varepsilon_1$  if necessary, it is no restriction to assume that

$$(7.3) \quad \eta_{\mathcal{O},1}(E_1(\varepsilon_1) \setminus E) < \frac{1}{10} \eta_{\mathcal{O},1}(E).$$

That may be seen as follows. Since  $\eta_{\mathcal{O},1}(E_1(\alpha))$  converges to  $\eta_{\mathcal{O},1}(E)$  as  $\alpha \rightarrow 0$ , and the limit is assumed to be positive, we have

$$(7.4) \quad \eta_{\mathcal{O},1}(E_1(\alpha) \setminus E) < \frac{1}{10} \eta_{\mathcal{O},1}(E)$$

for every small  $\alpha > 0$ . By Remark 7.3,

$$(7.5) \quad E_1(\alpha) \text{ is contained in the } \nu_k^{(n)}\text{-core of } E_1(\varepsilon_1)$$

for every large  $k$  and every small  $\alpha > 0$ . Clearly,

$$(7.6) \quad \eta_{\mathcal{O},1}(\partial E_1(\alpha)) = 0$$

for all but countably many values of  $\alpha$ . Fix  $\alpha > 0$ , depending only on  $\nu_{\mathcal{O}}$  and  $\varepsilon_1$ , satisfying (7.4), (7.5) and (7.6).

Applying Proposition 6.25 with  $X = E_1(\varepsilon_1)$ ,  $U = E_1(\alpha)$ ,  $\nu = \nu_k^{(n)}$  and  $\mathcal{T} = \mathcal{T}_{k,1}$ , for  $k$  large, we get a continuous Markov operator

$$\mathcal{T}'_{k,1} : \mathcal{B}(E_1(\alpha)) \rightarrow \mathcal{B}(E_1(\alpha))$$

adapted to  $(\nu_k^{(n)}, E_1(\alpha))$  and leaving invariant the normalized restriction  $\eta'_{k,1}$  of  $\eta_{k,1}$  to  $E(\alpha)$ . Replace  $\varepsilon_1$ ,  $\mathcal{T}_{k,1}$ , and  $\eta_{k,1}$  with  $\alpha$ ,  $\mathcal{T}'_{k,1}$ , and  $\eta'_{k,1}$ , respectively. Conditions (i) and (ii) in Theorem 7.1 are not affected by this: in particular, observe that (7.6) ensures that  $(\eta'_{k,1})_k \rightarrow \eta'_{\mathcal{O},1}$ . Then (7.3) just corresponds to (7.4).

We also introduce a suitable continuous self-coupling

$$(7.7) \quad \begin{aligned} & \tilde{\mathcal{T}}_{k,1} : \mathcal{B}(E_1(\varepsilon_1)^2) \rightarrow \mathcal{B}(E_1(\varepsilon_1)^2) \\ & \tilde{\mathcal{T}}_{k,1}\tilde{\varphi}(x, x') = \int_{E_1(\varepsilon_1)^2} \tilde{\varphi}(y, y') d\tilde{\sigma}_{k,1,x,x'}(y, y') \end{aligned}$$



of the operator  $\tilde{\mathcal{T}}_{k,1}$  such that

$$(7.8) \quad \tilde{\mathcal{T}}_{k,1}\tilde{\varphi}(x, x') = \int_G \tilde{\varphi}(gx, gx') d\nu_k^{(n)}(g)$$

when  $x$  and  $x'$  are both close to the equator  $E$ .

**2.** Consider  $\varepsilon_1''' \in (0, \varepsilon_1)$  and a sequence  $(\omega_{k,1})_k$  decreasing to zero. Property (7.3) ensures that

$$(7.9) \quad \eta_{k,1}(E_1(\varepsilon_1, \varepsilon_1''')) < \frac{2}{10}\eta_{k,1}(E_1(\varepsilon_1))$$

for every large  $k$ . Define  $A_k$  and  $B_k = B' \cup B_k''$  through

$$(7.10) \quad A_k = \{(x, x') \in E_1(\varepsilon_1''')^2 : d(x + x', E) > \omega_{k,1}\}$$

$$(7.11) \quad B' = \{(x, x') \in E_1(\varepsilon_1)^2 : d(x, E) > \varepsilon_1''' \text{ or } d(x', E) > \varepsilon_1'''\}$$

$$(7.12) \quad B_k'' = \{(x, x') \in E_1(\varepsilon_1)^2 : d(x + x', E) \leq \omega_{k,1}\}.$$

It is clear that  $A_k$  and  $B_k = B' \cup B_k''$  are disjoint, and their union is the whole  $E_1(\varepsilon_1)^2$ . For every large  $k$ , we have  $\omega_{k,1} < \varepsilon_1'''$  and, in view of the definition (7.1), that ensures that  $B'$  and  $B_k''$  are also disjoint. Moreover, (7.2) implies that  $B_k'' = \emptyset$  if  $\dim E = 1$ .

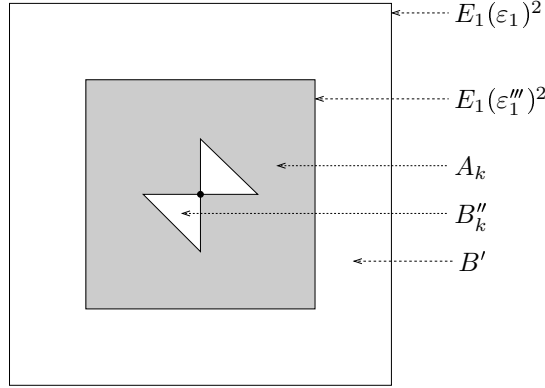


FIGURE 2. A sketch of the partition of  $E_1(\varepsilon_1)^2$  into the sets  $A_k$ ,  $B'$ , and  $B_k''$ . The latter converges to the central point  $(E, E)$  when  $k \rightarrow \infty$ , because  $\omega_{k,1} \rightarrow 0$ .

Suppose that for  $n \in \mathbb{N}$  sufficiently large there exist constants  $\kappa_A, \kappa_B > 0$  with  $\kappa_A > 9\kappa_B$ , and for each large  $k \in \mathbb{N}$  there exists a lower semi-continuous function  $\Psi_{k,1} : E_1(\varepsilon_1)^2 \rightarrow \mathbb{R}$  such that

$$(7.13) \quad \tilde{\mathcal{T}}_{k,1}\Psi_{k,1}(x, x') \leq \Psi_{k,1}(x, x') - \kappa_A \text{ for every } (x, x') \in A_k$$

$$(7.14) \quad \tilde{\mathcal{T}}_{k,1}\Psi_{k,1}(x, x') \leq \Psi_{k,1}(x, x') + \kappa_B \text{ for every } (x, x') \in B_k.$$

Then  $\Psi_{k,1}$  is a Margulis function for the operator  $\tilde{\mathcal{T}}_{k,1}$  relative to the partition  $(A_k, B_k)$  of  $E_1(\varepsilon_1)^2$ . Combining Proposition 6.23 with Lemma 6.21, we conclude

that there exist self-couplings  $\tilde{\eta}_{k,1,j}$  of  $\eta_{k,1}$  such that  $\int_{E_1(\varepsilon_1)^2} \Psi_{k,1} d\tilde{\eta}_{k,1,j}$  is finite and

$$(7.15) \quad \begin{aligned} \tilde{\eta}_{k,1,j}(B_k) &\geq \frac{\kappa_A}{\kappa_A + \kappa_B} \tilde{\eta}_{k,1,j}(E_1(\varepsilon_1)^2) \\ &> \frac{9}{10} \tilde{\eta}_{k,1,j}(E_1(\varepsilon_1)^2) = \frac{9}{10} \eta_{k,1}(E_1(\varepsilon_1)). \end{aligned}$$

Moreover,  $(\tilde{\eta}_{k,1,j})_j$  may be taken to converge to a  $\tilde{\mathcal{T}}_{k,1}$ -invariant self-coupling  $\tilde{\eta}_{k,1}$  of  $\eta_{k,1}$ .

The fact that  $\tilde{\eta}_{k,1,j}$  is a self-coupling of  $\eta_{k,1}$  together with the inequality (7.9) ensure that

$$(7.16) \quad \tilde{\eta}_{k,1,j}(B') \leq 2\eta_{k,1}(E_1(\varepsilon_1, \varepsilon_1'')) < \frac{4}{10} \eta_{k,1}(E_1(\varepsilon_1)).$$

Subtracting (7.16) from (7.15), we conclude that

$$(7.17) \quad \tilde{\eta}_{k,1,j}(B_k'') > \frac{5}{10} \eta_{k,1}(E_1(\varepsilon_1)) > 0.$$

This yields a contradiction when  $\dim E = 1$  because, as observed previously, in that case the set  $B_k''$  is empty.

**3.** If  $\dim E \geq 2$ , consider the map

$$\Sigma : E_1(\varepsilon_1)^2 \rightarrow \text{Gr}(2, d), \quad \Sigma(x, x') = x + x'.$$

The fact that  $\Sigma$  is not defined on the diagonal of  $E_1(\varepsilon_1)^2$  need not concern us at this stage: we will deal with it in Section 11.3. For each large  $k$ , let  $\eta_{k,2} = \Sigma_* \tilde{\eta}_{k,1}$  and  $\{\hat{\eta}_{k,1,y} : y \in \text{Gr}(2, d)\}$  be a disintegration of  $\tilde{\eta}_{k,1}$  with respect to the partition  $\{\Sigma^{-1}y : y \in \text{Gr}(2, d)\}$  of  $E_1(\varepsilon_1)^2$ . Then define

$$\mathcal{T}_{k,2} : \mathcal{B}(\Sigma(E_1(\varepsilon_1)^2)) \rightarrow \mathcal{B}(\Sigma(E_1(\varepsilon_1)^2)), \quad \mathcal{T}_{k,2}\psi(y) = \int_{\Sigma^{-1}y} \tilde{\mathcal{T}}_{k,1}(\psi \circ \Sigma) d\hat{\eta}_{k,1,y}.$$

The measure  $\eta_{k,2}$  is  $\mathcal{T}_{k,2}$ -invariant. Indeed, since  $\tilde{\eta}_{k,1}$  is  $\tilde{\mathcal{T}}_{k,1}$ -invariant,

$$\begin{aligned} \int_{\Sigma(E_1(\varepsilon_1)^2)} \mathcal{T}_{k,2}\psi d\eta_{k,2} &= \int_{\Sigma(E_1(\varepsilon_1)^2)} \int_{\Sigma^{-1}y} \tilde{\mathcal{T}}_{k,1}(\psi \circ \Sigma) d\hat{\eta}_{k,1,y} d\eta_{k,2}(y) \\ &= \int_{E_1(\varepsilon_1)^2} \tilde{\mathcal{T}}_{k,1}(\psi \circ \Sigma) d\tilde{\eta}_{k,1} = \int_{E_1(\varepsilon_1)^2} (\psi \circ \Sigma) d\tilde{\eta}_{k,1} \\ &= \int_{\Sigma(E_1(\varepsilon_1)^2)} \psi d\eta_{k,2} \end{aligned}$$

for any  $\psi \in \mathcal{B}(\Sigma(E_1(\varepsilon_1)^2))$ . Taking the limit as  $j \rightarrow \infty$  in (7.17), we find that

$$(7.18) \quad \tilde{\eta}_{k,1}(B_k'') \geq \frac{5}{10} \eta_{k,1}(E_1(\varepsilon_1)).$$

Recall that we take  $\omega_{k,1} \rightarrow 0$  as  $k \rightarrow \infty$ . Then, in view of the definition of  $B_k''$  in (7.12), the sets  $\Sigma(B_k'')$  approach  $E_2$ . Taking the limit  $k \rightarrow \infty$  in (7.18), we find that

$$\eta_{\infty,2}(E_2) \geq \frac{5}{10} \eta_{\infty,1}(E_1(\varepsilon_1)) > 0.$$

Now define  $n_2 = n$ . If  $y$  is close to  $E_2$  in  $\text{Gr}(2, d)$  then  $x$  and  $x'$  are close to  $E_1$  in  $\text{Gr}(1, d)$  and, in particular, they are in the  $\nu_k^{(n_2)}$ -core of  $E_1(\varepsilon_1)$  for every large  $k$ .

Then, using (7.8),

$$\begin{aligned}\mathcal{T}_{k,2}\psi(y) &= \int_{\Sigma^{-1}y} \int_G \psi \circ \Sigma(gx, gx') d\nu_k^{(n_2)}(g) d\tilde{\eta}_{k,1,y}(x, x') \\ &= \int_{\Sigma^{-1}y} \int_G \psi(gy) d\nu_k^{(n_2)}(g) d\tilde{\eta}_{k,1,y}(x, x') = \int_G \psi(gy) d\nu_k^{(n_2)}(g),\end{aligned}$$

because each  $\tilde{\eta}_{k,1,y}$  is a probability. Pick  $\varepsilon_2 > 0$  such that this holds for every  $y \in E_2(\varepsilon_2)$ . Apply Proposition 6.25 with  $X = \Sigma(E_1(\varepsilon_1)^2)$ ,  $U = E_2(\varepsilon_2)$ ,  $\nu = \nu_k^{(n_2)}$ , and  $\mathcal{T} = \mathcal{T}_{k,2}$ , for  $k$  large. Replace  $\mathcal{T}_{k,2}$  and  $\eta_{k,2}$  with this new Markov operator and invariant measure, respectively.

This would complete the proof of Theorem 7.1 for  $r = 1$ .

4. However, in general we cannot construct a Margulis function  $\Psi_{k,1}$  as required. Essentially, the problem is that, since we do not control the measures  $\tilde{\sigma}_{k,1,x,x'}$  in the border region, the inequality (7.14) cannot be proved to hold unless the Margulis function  $\Psi_{k,1}$  is taken to be bounded in the border region. However, outside the border region  $\Psi_{k,1}(x, x')$  must be very large if  $x$  and  $x'$  are close. The only way to reconcile these two requirements is to introduce some drastic discontinuities in  $\Psi_{k,1}$  which then cause (7.14) to fail at some points.

This problem is solved as follows. We do indeed create a discontinuity by cutting  $\Psi_{k,1}$  off in such a way that it is bounded in the border region. The main idea for dealing with the discontinuity, that we call *recoupling*, involves replacing  $\tilde{\mathcal{T}}_{k,1}$  with another Markov operator  $\hat{\mathcal{T}}_{k,1}$  that still projects to  $\mathcal{T}_{k,1}$  on either factor. The recoupling modification is restricted to a region which is disjoint from  $B_k''$  and relatively far from the equator, so that the properties of the measures  $\eta_{k,2}$  are not affected. These arguments are detailed in Sections 8 through 11.

In Section 8 we introduce the notion of *vertical projection*  $\text{VP}_1(x, x')$  of a pair of points  $x$  and  $x'$  in  $P$ , and we use it to construct a candidate  $-\log \text{VP}_1(x, x')$  to a Margulis function for  $\tilde{\mathcal{T}}_{\infty,1}$ . The problem with this function is that it refers explicitly to  $E$  and, since the equator is not  $\nu_k$ -invariant for  $k \in \mathbb{N}$ , the estimates break down for  $\tilde{\mathcal{T}}_{k,1}$  when  $k$  is finite.

This is fixed in Section 9.1, where we replace the vertical projection with a kind of cut-off that we call the *stabilized vertical projection* and denote as  $\text{SVP}_1(x, x')$ . Then, in Section 9.2 we give the precise recipe for the other cut-off, that was mentioned previously. By the end of Section 10 point we will have a much more viable candidate  $\Psi_{k,1}$  for Margulis function when  $k$  is finite.

The details of the recoupling procedure are described in Section 11.1. Proposition 11.2 essentially states that this function  $\Psi_{k,1}$  is a Margulis function for the recoupled Markov operator  $\hat{\mathcal{T}}_{k,1}$ . In Sections 11.2 and 11.3 we wrap the arguments up to conclude the step  $r = 1$  of the proof.

**7.3. General step.** For  $r > 1$  the strategy is similar, except that we need to work with *flag varieties*. Recall that  $\mathcal{F}(r, d)$  denotes the space of flags

$$F_1 \subset F_2 \subset \cdots \subset F_{r-1} \subset F_r \subset \mathbb{R}^d,$$

where each  $F_i$  has dimension  $i$ . We use  $x = (F_1, \dots, F_r)$  and  $x' = (F'_1, \dots, F'_r)$  to denote generic elements of  $\mathcal{F}(r, d)$ . Define  $E_r^\diamond = \{x \in \mathcal{F}(r, d) : F_r \in E_r\}$  and, for each  $\varepsilon > \varepsilon' > 0$ ,

$$E_r^\diamond(\varepsilon) = \{x \in \mathcal{F}(r, d) : F_r \in E_r(\varepsilon)\} \text{ and } E_r^\diamond(\varepsilon, \varepsilon') = E_r^\diamond(\varepsilon) \setminus E_r^\diamond(\varepsilon').$$

By definition, for any  $\nu \in \mathcal{P}_c(G)$ , the  $\nu$ -core of  $E^\circ(\varepsilon)$  is the subset of flags  $x$  such that  $F_r \in \mathcal{X}_\nu E_r(\varepsilon_r)$ , and the  $\nu$ -border of  $E^\circ(\varepsilon)$  is defined in a similar fashion.

We consider  $E_r^\circ(\varepsilon_r)^2 = E_r^\circ(\varepsilon_r) \times E_r^\circ(\varepsilon_r)$  and we aim to construct a Margulis function  $\Psi_{k,r}$  for a suitable Markov operator  $\widehat{\mathcal{T}}_{k,r}^\circ : \mathcal{B}(E_r^\circ(\varepsilon_r)^2) \rightarrow \mathcal{B}(E_r^\circ(\varepsilon_r)^2)$  relative to a convenient generalization  $E_r^\circ(\varepsilon_r)^2 = A_k \cup B_k' \cup B_k''$  of the partition (7.10)–(7.12). For  $i = 1, 2$ , define  $\theta_i : E_r^\circ(\varepsilon_r)^2 \rightarrow \text{Gr}(r, d)$  to be the projection to either factor followed by the canonical map

$$(7.19) \quad \mathcal{F}(r, d) \rightarrow \text{Gr}(r, d), \quad x \mapsto F_r.$$

A constraint on  $\widehat{\mathcal{T}}_{k,r}^\circ$  is that it must project to  $\mathcal{T}_{k,r}$  under both  $\theta_i$ . Then we can find a  $\widehat{\mathcal{T}}_{k,r}^\circ$ -invariant measure  $\widehat{\eta}_{k,r}^\circ$  which maps to  $\eta_{k,r}$  under both projections.

As in the case  $r = 1$ , once we have constructed  $\widehat{\mathcal{T}}_{k,r}^\circ$ ,  $\Psi_{k,r}$ , and  $\widehat{\eta}_{k,r}^\circ$  we can try to get an estimate of the form

$$\widehat{\eta}_{k,r}^\circ(B_k'') > c \widehat{\eta}_{k,r}^\circ(E_r^\circ(\varepsilon_r)^2),$$

where  $c > 0$  is some absolute constant. Then, if we push  $\widehat{\eta}_{k,r}^\circ$  forward by one of the maps

$$\begin{aligned} \Sigma_r : E_r^\circ(\varepsilon_r)^2 &\rightarrow \text{Gr}(r+1, d), & \Sigma_r(x, x') &= F_1' + F_r \\ \Sigma_r : E_r^\circ(\varepsilon_r)^2 &\rightarrow \text{Gr}(r+1, d), & \Sigma_r(x, x') &= F_1 + F_r' \end{aligned}$$

we obtain probability measures  $\eta_{k,r+1}$  on  $\text{Gr}(r+1, d)$  satisfying the conditions of Theorem 7.1.

However, for  $r > 1$  the simple cut-off procedure we use in the initial step of the induction is no longer sufficient to ensure that the Margulis function is bounded on the border region. To fix that, we take advantage of the additional freedom provided by the flag space, which is that we may vary the projections to  $\text{Gr}(j, d)$  for  $j < r$ . More precisely, we modify the dynamics by averaging ("spreading out") the Markov operator over the subspaces of dimension less than  $r$  in the flag. Thus we end up with modified Markov operators  $\widehat{\mathcal{T}}_{k,r}^\circ$  that still project to the  $\mathcal{T}_{k,r}$  under both  $\theta_i$ . Then the kind of argument we sketched in the previous paragraphs can actually be carried out for  $\widehat{\mathcal{T}}_{k,r}^\circ$ -invariant measures  $\widehat{\eta}_{k,r}^\circ$  that map to  $\eta_{k,r}$  under both projections.

The detailed arguments are structured as follows in Sections 12 through 16. In Section 12 we extend the notions of *vertical angle* and *vertical projection* to  $r > 1$ , and we use them to exhibit a candidate  $-\log \psi_r(x, x')$  to a Margulis function. In Section 13.1 we move to introduce the  $r > 1$  version of the *stabilized vertical projection*  $\text{SVP}_r(x, x')$ , and in Section 13.2 we describe the corresponding version for the cut-off procedure. In Section 14 we extend this analysis to the perturbed random walks, that is, to finite  $k$ . The Margulis function  $\Psi_{k,r}$  is defined at the end of that section.

The spreading out modification is detailed in Section 15. Essentially, we define  $\widehat{\mathcal{T}}_{k,r}^\circ = \widehat{\mathcal{T}}_{q,r}^\circ \circ \widetilde{\mathcal{Q}}_r$  where  $\widetilde{\mathcal{Q}}_r : Bd(E_r^\circ(\varepsilon_r)^2) \rightarrow \mathcal{B}(E_r^\circ(\varepsilon_r)^2)$  has the form

$$\widetilde{\mathcal{Q}}_r \psi(x, x') = \int \Psi d\tilde{q}_{r,x,x'}$$

where  $\tilde{q}_{r,x,x'}$  is a smooth probability measure on the subset of pairs of flags whose  $r$ -coordinate is  $(F_r, F_r')$ . A relevant feature is that  $\widetilde{\mathcal{Q}}_r$  maps to the identity under the canonical map (7.19). The details of the recoupling procedure are described in Section 16.2. In Sections 16.3 and 16.4 we wrap up the proof.

### Part III. First step of the induction

#### 8. PREPARING A MARGULIS FUNCTION

We are going to construct a positive function  $\psi_1$  such that  $-\log \psi_1$  has some features of a Margulis function for large iterates of the operator  $\mathcal{P}_{\nu_\infty}$ :

**Proposition 8.1.** *There exist  $\kappa'_1 = \kappa'_1(\nu_\infty) > 0$  and  $C'_1 = C'_1(\nu_\infty) > 0$  and for any  $\delta > 0$  there exists  $N_1 = N_1(\nu_\infty, \delta) \in \mathbb{N}$  such that for every  $n \geq N_1$  there exists  $\rho'_1 = \rho'_1(\nu_\infty, \delta, n) > 0$  satisfying*

$$(8.1) \quad \int_G -\log \psi_1(gx, gx') d\nu_\infty^{(n)}(g) \leq -\log \psi_1(x, x') - (\kappa'_1 - C'_1 \delta)n$$

for every  $x \neq x'$  in  $E_1(\rho'_1)$ .

The conclusion of Proposition 8.1 does not quite fit the definition of a Margulis function because the set  $E_1(\rho'_1) \times E_1(\rho'_1)$  where the estimate holds is not invariant under the  $G$ -action restricted to  $\text{supp } \nu_\infty$  and thus  $\mathcal{P}_{\nu_\infty}$  cannot be considered a Markov operator on this set. Nevertheless,  $-\log \psi_1$  is an important ingredient in the definition of the actual Margulis function, which will be completed in Section 11.

The proof of Proposition 8.1 occupies the remainder of this section. Throughout, we think of the equator as being “horizontal” and use the word “vertical” to refer to the orthogonal direction. The numbers  $\delta > 0$  and  $n \in \mathbb{N}$  will remain as free parameters for most of our arguments. Near the end, in (11.13), we will fix  $\delta > 0$  to be small, depending only on  $\nu_\infty$ , and  $n \in \mathbb{N}$  large, depending on  $\nu_\infty$  and  $\delta > 0$ .

Fix  $A = A(\nu_\infty) > 0$  and a compact neighborhood  $\mathcal{W}_0 = \mathcal{W}_0(\nu_\infty) \subset G$  of the support of  $\nu_\infty$  such that

$$-\log d(U, V) - A \leq -\log d(fU, fV) \leq -\log d(U, V) + A$$

for any  $f \in \mathcal{W}_0$ , any pair  $U \neq V$  in  $\text{Gr}(r, d)$  and any  $1 \leq r \leq d$ . Moreover, let  $B = B(\nu_\infty) > 0$  be defined by

$$B = \sup \{ \log \|f\| + \log \|f^{-1}\| : f \in \mathcal{W}_0 \} + 2.$$

Since  $(\nu_k)_k$  converges to  $\nu_\infty$  in the space  $\mathcal{P}_c(G)$ , it is no restriction to assume that  $\text{supp } \nu_k \subset \mathcal{W}_0$  for every  $k \in \mathbb{N}$ . In particular,

$$(8.2) \quad -\log d(fU, fV) \leq -\log d(U, V) + An$$

$$(8.3) \quad \log \|f\| \|f^{-1}\| \leq Bn$$

for every  $U \neq V$  in  $\text{Gr}(r, d)$ ,  $1 \leq r \leq d$ ,  $f \in \text{supp } \nu_k^{(n)}$ ,  $k \in \mathbb{N} \cup \{\infty\}$ , and  $n \in \mathbb{N}$ .

**8.1. Vertical angle function.** Given distinct points  $x$  and  $x'$  in  $P = \text{Gr}(1, d)$ , let  $y = y(x, x') \subset P$  denote the *great circle* through  $x$  and  $x'$ , that is, the subset of  $P$  associated with the element  $x + x'$  of  $\text{Gr}(2, d)$  that contains both  $x$  and  $x'$ .

The *vertical angle function*  $\text{VA}_1$  is the sine of the angle between the great circle  $y = y(x, x')$  and the equator, that is (recall (7.1) and check Figure 3),

$$(8.4) \quad \text{VA}_1(x, x') = d(y, E) = \sup_{u \in y} d(u, E) = \sup_{u \in y} \frac{\|u^\perp\|}{\|u\|}.$$

For any  $u \in y$  that realizes the supremum,

$$(8.5) \quad \text{VA}_1(gx, gx') \geq \frac{\|(gu)^\perp\|}{\|gu\|} = \text{VA}_1(x, x') \frac{\|(gu)^\perp\|}{\|gu\|} \frac{\|u\|}{\|u^\perp\|}$$

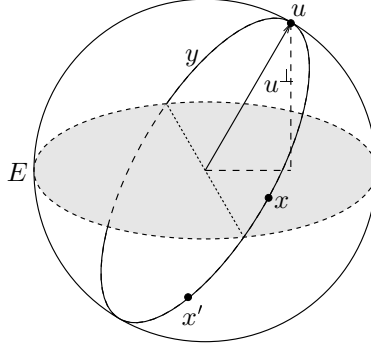


FIGURE 3. Geometric meaning of the vertical angle function  $\text{VA}_1$ . By making  $x$  and  $x'$  go to  $E$  along a given great circle, one sees that  $\text{VA}_1(x, x')$  need not go to infinity when  $x$  and  $x'$  approach the equator.

for any  $g \in G$ . By Remark 3.1, when  $g \in \text{supp } \nu_\infty^{(n)}$  this means that

$$(8.6) \quad \text{VA}_1(gx, gx') \geq \frac{\|g^\perp u^\perp\|}{\|gu\|} = \text{VA}_1(x, x') \frac{\|g^\perp u^\perp\|}{\|gu\|} \frac{\|u\|}{\|u^\perp\|}.$$

Still for  $g \in \text{supp } \nu_\infty^{(n)}$ , using Remark 3.1 and (8.3) we get that

$$\frac{\|g^\perp u^\perp\|}{\|gu\|} \frac{\|u\|}{\|u^\perp\|} \geq \frac{1}{\|(g^\perp)^{-1}\| \|g\|} \geq \frac{1}{\|g^{-1}\| \|g\|} \geq e^{-Bn},$$

and so

$$(8.7) \quad -\log \text{VA}_1(gx, gx') \leq -\log \text{VA}_1(x, x') + Bn.$$

**Lemma 8.2.** *There exists  $\tilde{\kappa}_1 = \tilde{\kappa}_1(\nu_\infty) > 0$  and for each  $\delta > 0$  there exist  $\tilde{\theta}_1 = \tilde{\theta}_1(\nu_\infty, \delta) > 0$  and  $\tilde{N}_1 = \tilde{N}_1(\nu_\infty, \delta) \in \mathbb{N}$  such that for every  $n \geq \tilde{N}_1$  and  $x \neq x'$  in  $P$  there exists  $\tilde{\mathcal{E}}_1 = \tilde{\mathcal{E}}_1(\nu_\infty, \delta, n, x, x') \subset \text{supp } \nu_\infty^{(n)}$  with  $\nu_\infty^{(n)}(\tilde{\mathcal{E}}_1^c) < \delta$  and*

$$(8.8) \quad -\log \text{VA}_1(gx, gx') \leq \max\{-\log \text{VA}_1(x, x') - \tilde{\kappa}_1 n, \tilde{\theta}_1\} \text{ for every } g \in \tilde{\mathcal{E}}_1.$$

*Proof.* Let  $\kappa_0 = \kappa_0(\nu_\infty) > 0$ ,  $N_0 = N_0(\nu_\infty, \delta) \in \mathbb{N}$ ,  $\mathcal{E}_0 = \mathcal{E}_0(\nu_\infty, \delta, n, u^\perp) \subset \text{supp } \nu_\infty^{(n)}$ , and  $\tau_0 = \tau_0(\nu_\infty, \delta) > 0$  be as in Proposition 5.5. Given  $x \neq x'$  in  $P$ , take  $u \in y$  realizing the supremum in (8.4). Write  $u = u^E + u^\perp$  with  $u^E \in E$  and  $u^\perp \in E^\perp$ . Take

$$(8.9) \quad \begin{aligned} \tilde{\kappa}_1 &= \kappa_0/2, \quad \tilde{\theta}_1 = -\log(\tau_0/4), \\ \tilde{N}_1 &> \max\{N_0, 4/\tilde{\kappa}_1\}, \quad \text{and } \tilde{\mathcal{E}}_1 = \mathcal{E}_0(\nu_\infty, \delta, n, u^\perp). \end{aligned}$$

Let  $n \geq \tilde{N}_1$  and  $g \in \tilde{\mathcal{E}}_1 \subset \text{supp } \nu_\infty^{(n)}$ . If  $\|g^\perp u^\perp\|/\|gu\| \geq \tau_0/2$  then the inequality in (8.6) implies that

$$(8.10) \quad -\log \text{VA}_1(gx, gx') \leq -\log \frac{\tau_0}{2} \leq \tilde{\theta}_1.$$

If  $\|g^\perp u^\perp\|/\|gu\| < \tau_0/2$  then part (2) of Proposition 5.5 gives that

$$\frac{\|gu^\perp\|}{\|gu\|} < \frac{1}{2}, \text{ which implies } \frac{\|gu^E\|}{\|gu\|} > \frac{1}{2}.$$

Substituting the latter inequality and  $\|u\| \geq \|u^E\|$  in (8.6), we find that

$$(8.11) \quad \text{VA}_1(gx, gx') \geq \frac{1}{2} \text{VA}_1(x, x') \frac{\|g^\perp u^\perp\|}{\|u^\perp\|} \frac{\|u^E\|}{\|gu^E\|}.$$

Thus, recalling the definition (5.3) and part (1) of Proposition 5.5,

$$\text{VA}_1(gx, gx') \geq \frac{1}{2} \text{VA}_1(x, x') \frac{\|Dg_{u^E}^\perp u^\perp\|}{\|u^\perp\|} \geq \frac{1}{2} \text{VA}_1(x, x') e^{\kappa_0 n}.$$

By the choices of  $\tilde{\kappa}_1$  and  $\tilde{N}_1$  in (8.9), this implies that

$$(8.12) \quad \begin{aligned} -\log \text{VA}_1(gx, gx') &\leq -\log \text{VA}_1(x, x') + \log 2 - 2\tilde{\kappa}_1 n \\ &\leq -\log \text{VA}_1(x, x') - \tilde{\kappa}_1 n. \end{aligned}$$

The conclusion of the lemma is contained in (8.10) and (8.12).  $\square$

**8.2. Vertical projection function.** The function  $-\log \text{VA}_1(x, x')$  cannot be used as a Margulis function to detect the equator because (unless  $\dim E = 1$ ) it is possible that  $-\log \text{VA}_1(x, x')$  remains bounded even as  $d(x, E)$  and  $d(x', E)$  go to zero: the great circle through points close to the equator need not be close to the equator, as illustrated in Figure 3.

To (partially) rectify this, we introduce the *vertical projection* function  $\text{VP}_1$ , defined as follows:

$$(8.13) \quad \text{VP}_1(x, x') = \text{VA}_1(x, x') d(x, x')^{\gamma_1}$$

where  $\gamma_1 = \gamma_1(\nu_\infty)$  is a small positive constant to be chosen shortly (Proposition 8.3). Note that if  $\gamma_1 = 1$  then  $\text{VP}_1(x, x')$  would indeed be a sort of projection on the orthogonal complement to the equator. It is clear from the definitions (8.4) and (8.13) that  $\text{VA}_1$  and  $\text{VP}_1$  are symmetric functions:

$$(8.14) \quad \text{VA}_1(x, x') = \text{VA}_1(x', x) \text{ and } \text{VP}_1(x, x') = \text{VP}_1(x', x) \text{ for any } x \neq x'.$$

**Proposition 8.3.** *There exist  $\gamma_1 = \gamma_1(\nu_\infty) > 0$  and  $\kappa'_1 = \kappa'_1(\nu_\infty) > 0$  and for each  $\delta > 0$  there exists  $N_1 = N_1(\nu_\infty, \delta) \in \mathbb{N}$  such that for every  $n \geq N_1$  there exists  $\rho'_1 = \rho'_1(\nu_\infty, \delta, n) > 0$  such that for any  $x \neq x'$  in  $E_1(\rho'_1)$  there exists  $\mathcal{E}'_1 = \mathcal{E}'_1(\nu_\infty, \delta, n, x, x') \subset \text{supp } \nu_\infty^{(n)}$  with  $\nu_\infty^{(n)}((\mathcal{E}'_1)^c) < \delta$  and*

$$(8.15) \quad -\log \text{VP}_1(gx, gx') \leq -\log \text{VP}_1(x, x') - \kappa'_1 n \text{ for every } g \in \mathcal{E}'_1.$$

*Proof.* The overall strategy goes as follows. If the vertical angle  $\text{VA}_1(x, x')$  is small then the conclusion of the present proposition is a consequence of Lemma 8.2, provided that we choose the constant  $\gamma_1$  small enough to make the variation of  $d(x, x')^{\gamma_1}$  negligible. If  $\text{VA}_1(x, x')$  is large, let  $v$  and  $v'$  be unit vectors in the direction of  $x$  and  $x'$  and such that the angle between them is non-obtuse (the latter may always be obtained by replacing  $v$  with  $-v$  if necessary). In Lemma 8.4 we check that if  $x$  is close to  $E$  then the difference  $w = v' - v$  is roughly vertical, meaning that the angle between  $w$  and the equator is bounded from below. See Figure 4. In Lemma 8.5 we deduce that in this situation  $-\log d(x, x')$  decreases under most large iterates: the reason is that the dynamics increases vertical components and, for  $n$  sufficiently large, the vertical component of the iterate of  $w$  dominates. This implies the conclusion of the proposition because in this regime the variation of  $\text{VA}_1(x, x')$  is bounded. Let us fill-in the details.

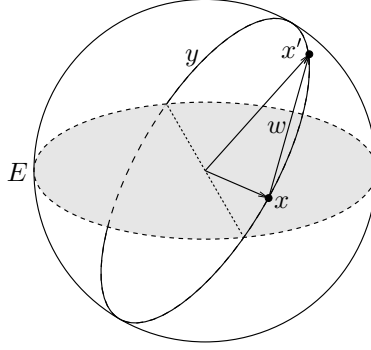


FIGURE 4. Interpreting Lemma 8.4: if  $d(x, E)$  is small compared to  $\text{VA}_1(x, x') = d(y, E)$  then  $w = v' - v$  is roughly vertical.

**Lemma 8.4.** *Given  $x \neq x'$  in  $P$ , let  $w = v' - v$  be the difference between unit vectors in the directions of  $x$  and  $x'$ , respectively. Then*

$$(8.16) \quad d(x, E) < \frac{1}{4} \text{VA}_1(x, x') \text{ implies } \frac{\|w^\perp\|}{\|w\|} > \frac{1}{4} \text{VA}_1(x, x').$$

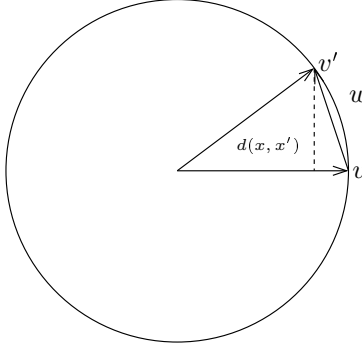


FIGURE 5. Illustrating the estimates in (8.17): the distance from  $x$  to  $x'$  is given by the length of the vertical dashed segment, which is bounded above by  $\|w\|$ ; the angle between  $v$  and  $w$  is at least  $\pi/4$ , as the angle between  $v$  and  $v'$  is non-obtuse.

*Proof.* As the angle between  $v$  and  $v'$  was taken to be non-obtuse (check Figure 5)

$$(8.17) \quad d(x, x') = |\sin \angle(v, v')| \leq \|w\| \text{ and } \angle(w, v) \geq \frac{\pi}{4}.$$

Take  $u \in y$  realizing the supremum in (8.4). Since  $y = \text{span}\{v, v'\} = \text{span}\{v, w\}$ , we may write  $u = av + bw$  with  $a, b \in \mathbb{R}$ . The angle bound in (8.17) implies that  $\|av\|$



and  $\|bw\|$  are both less than  $2\|u\|$ . Thus,

$$\begin{aligned} \text{VA}_1(x, x') &= \frac{\|u^\perp\|}{\|u\|} \leq \frac{\|av^\perp\|}{\|u\|} + \frac{\|bw^\perp\|}{\|u\|} \\ &< 2\frac{\|v^\perp\|}{\|v\|} + 2\frac{\|w^\perp\|}{\|w\|} \leq 2d(x, E) + 2\frac{\|w^\perp\|}{\|w\|}. \end{aligned}$$

Thus,  $d(x, E)$  and  $\|w^\perp\|/\|w\|$  cannot be both less than  $\text{VA}_1(x, x')/4$ .  $\square$

Take  $\tilde{\kappa}_1 = \tilde{\kappa}_1(\nu_\infty) > 0$  and  $\tilde{\theta}_1 = \tilde{\theta}_1(\nu_\infty, \delta) > 0$  to be as in Lemma 8.2.

**Lemma 8.5.** *There exists  $\hat{\kappa}_1 = \hat{\kappa}_1(\nu_\infty) > 0$  and for each  $\delta > 0$  there exists  $\hat{N}_1 = \hat{N}_1(\nu_\infty, \delta) \in \mathbb{N}$  such that for each  $n \geq \hat{N}_1$  there exists  $\hat{\rho}_1 = \hat{\rho}_1(\nu_\infty, \delta, n) > 0$  such that for any  $x \neq x'$  in  $E_1(\hat{\rho}_1)$  with  $-\log \text{VA}_1(x, x') \leq \tilde{\theta}_1 + \tilde{\kappa}_1 n$  there exists  $\hat{\mathcal{E}}_1 = \hat{\mathcal{E}}_1(\nu_\infty, \delta, n, x, x') \subset \text{supp } \nu_\infty^{(n)}$  with  $\nu_\infty^{(n)}(\hat{\mathcal{E}}_1^c) < \delta$  and*

$$(8.18) \quad -\log d(gx, gx') \leq -\log d(x, x') - \log \text{VA}_1(x, x') - \hat{\kappa}_1 n \text{ for every } g \in \hat{\mathcal{E}}_1.$$

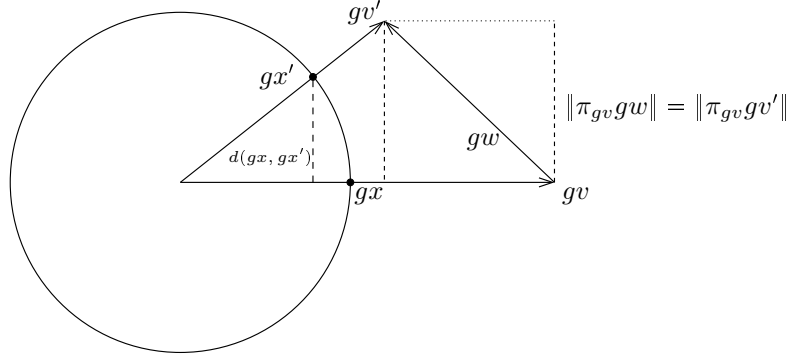


FIGURE 6. Verifying the inequality (8.21), when  $\|gv\| \geq \|gv'\|$ : the distance between  $gx$  and  $gx'$  is given by the length of the vertical dashed segment on the left, which is greater than the length of the vertical dashed segment on the right divided by  $\|gv\|$ .

*Proof.* Let  $\kappa_0 = \kappa_0(\nu_\infty) > 0$ ,  $N_0 = N_0(\nu_\infty, \delta) \in \mathbb{N}$ , and  $\mathcal{E}_0 = \mathcal{E}_0(\nu_\infty, \delta, n, v^\perp) \subset \text{supp } \nu_\infty^{(n)}$  be as in Proposition 5.5. Given  $x \neq x'$  in  $P$ , let  $w = v' - v$  be the difference between unit vectors  $v$  and  $v'$  in the directions of  $x$  and  $x'$ , respectively. Take

$$(8.19) \quad \begin{aligned} \hat{\kappa}_1 &= \kappa_0/2, \quad \hat{N}_1 > \max\{N_0, 5/\hat{\kappa}_1\}, \\ \hat{\rho}_1 &< e^{-\tilde{\theta}_1 - \tilde{\kappa}_1 n}/10, \text{ and } \hat{\mathcal{E}}_1 = \mathcal{E}_0(\nu_\infty, \delta, n, w^\perp). \end{aligned}$$

Let  $n \geq \hat{N}_1$  and  $g \in \hat{\mathcal{E}}_1 \subset \text{supp } \nu_\infty^{(n)}$ . As observed in (8.17),

$$(8.20) \quad d(x, x') \leq \|w\|$$

Let us suppose that  $\|gv\| \geq \|gv'\|$ ; the case  $\|gv\| \leq \|gv'\|$  is analogous, reversing the roles of  $x$  and  $x'$ . Then (check Figure 6)

$$(8.21) \quad d(gx, gx') = \left\| \Pi_{gv} \frac{gv'}{\|gv'\|} \right\| = \left\| \Pi_{gv} \frac{gw}{\|gv'\|} \right\| \geq \frac{1}{\|gv\|} \|\Pi_{gv} gw\|,$$

and so (keep in mind that  $\|v\| = \|v'\| = 1$ ),

$$(8.22) \quad -\log d(gx, gx') \leq -\log d(x, x') - \log \frac{\|\Pi_{gv}gw\|}{\|gv\|} \frac{\|v\|}{\|w\|}.$$

By the condition on  $\hat{\rho}_1$  in (8.19), if  $d(x, E) \leq \hat{\rho}_1$  then

$$4d(x, E) < e^{-\hat{\theta}_1 - \hat{\kappa}_1 n} < \text{VA}_1(x, x'),$$

and then Lemma 8.4 gives that

$$(8.23) \quad \frac{\|w^\perp\|}{\|w\|} > \frac{1}{4} \text{VA}_1(x, x') > \frac{1}{4} e^{-\hat{\theta}_1 - \hat{\kappa}_1 n}.$$

Up to further reducing  $\hat{\rho}_1$ , we may also assume that

$$(8.24) \quad \frac{\|hz\|}{\|z\|} \geq \frac{1}{2} \frac{\|hz^E\|}{\|z^E\|} \text{ and } \|\Pi_{hz}hw\| \geq \frac{1}{2} \|\Pi_{hz^E}hw\|$$

for any non-zero  $z = z^E + z^\perp$  in  $E \oplus E^\perp$  with  $\|z^\perp\|/\|z\| \leq \hat{\rho}_1$  and any  $h \in \text{supp } \nu_{\mathcal{O}}^{(n)}$ . Indeed, the first part of (8.24) is a simple consequence of continuity; in the second one note also that  $w$  is bounded away from the horizontal, by (8.23). This last part of (8.24) implies that

$$(8.25) \quad \|\Pi_{hz}hw\| \geq \frac{1}{2} \|\Pi_{hz^E}hw\| \geq \frac{1}{2} \|\Pi_Ehw\| = \frac{1}{2} \|(hw)^\perp\| = \frac{1}{2} \|h^\perp w^\perp\|.$$

Noting that  $\|v^\perp\|/\|v\| = d(x, E) \leq \hat{\rho}_1$ , take  $z = v$  and  $h = g$  in the previous two relations. Thus, substituting (8.24) and (8.25) in (8.22),

$$-\log d(gx, gx') \leq -\log d(x, x') + \log 4 - \log \frac{\|g^\perp w^\perp\|}{\|w\|} \frac{\|v^E\|}{\|gv^E\|}.$$

Then, using also (8.23) and (5.3),

$$\begin{aligned} -\log d(gx, gx') &\leq -\log d(x, x') - \log \text{VA}_1(x, x') + \log 16 - \log \frac{\|g^\perp w^\perp\|}{\|w^\perp\|} \frac{\|v^E\|}{\|gv^E\|} \\ &= -\log d(x, x') - \log \text{VA}_1(x, x') + \log 16 - \log \frac{\|Dg_{v^E}^\perp w^\perp\|}{\|w^\perp\|}. \end{aligned}$$

By part (1) of Proposition 5.5 and the choice of  $\hat{\kappa}_1$  and  $\hat{N}_1$  in (8.19), it follows that

$$\begin{aligned} -\log d(gx, gx') &\leq -\log d(x, x') - \log \text{VA}_1(x, x') + 5 - 2\hat{\kappa}_1 n \\ &\leq -\log d(x, x') - \log \text{VA}_1(x, x') - \hat{\kappa}_1 n, \end{aligned}$$

as claimed.  $\square$

We are ready to prove Proposition 8.3. Take  $A = A(\nu_{\mathcal{O}})$  as in (8.2) and then define

$$(8.26) \quad \begin{aligned} \gamma_1 &= \min \{1, \tilde{\kappa}_1/(2A)\}, \\ \kappa'_1 &= \min \{\tilde{\kappa}_1/2, \gamma_1 \hat{\kappa}_1/2\}, \\ \rho'_1 &= \min \{\hat{\rho}_1(\nu_{\mathcal{O}}, \delta), \hat{\rho}_1(\nu_{\mathcal{O}}, \delta/2)\} \\ N_1 &> \max \{\tilde{N}_1(\nu_{\mathcal{O}}, \delta/2), \hat{N}_1(\nu_{\mathcal{O}}, \delta/2), 2\tilde{\theta}_1/(\gamma \hat{\kappa}_1)\} \text{ and} \\ \mathcal{E}'_1 &= \tilde{\mathcal{E}}_1(\nu_{\mathcal{O}}, \delta/2, n, x, x') \cap \hat{\mathcal{E}}_1(\nu_{\mathcal{O}}, \delta/2, n, x, x'). \end{aligned}$$

Observe that  $\nu_{\mathcal{O}}^{(n)}((\mathcal{E}'_1)^c) < \delta$ . By definition,

$$(8.27) \quad -\log \text{VP}_1(gx, gx') = -\log \text{VA}_1(gx, gx') - \gamma_1 \log d(gx, gx').$$

Consider  $x \neq x'$  in  $E_1(\rho'_1)$ ,  $n \geq N_1$  and  $g \in \mathcal{E}'_1$ . If  $-\log \text{VA}_1(x, x') \geq \tilde{\theta}_1 + \tilde{\kappa}_1 n$  then, by Lemma 8.2,

$$(8.28) \quad -\log \text{VA}_1(gx, gx') \leq -\log \text{VA}_1(x, x') - \tilde{\kappa}_1 n$$

Substituting (8.2) and (8.28) in (8.27) we find that

$$(8.29) \quad \begin{aligned} -\log \text{VP}_1(gx, gx') &\leq -\log \text{VA}_1(x, x') - \tilde{\kappa}_1 n - \gamma_1 \log d(x, x') + \gamma_1 An \\ &\leq -\log \text{VP}_1(x, x') - \frac{\tilde{\kappa}_1}{2} n \leq -\log \text{VP}_1(x, x') - \kappa'_1 n. \end{aligned}$$

Now assume that  $-\log \text{VA}_1(x, x') \leq \tilde{\theta}_1 + \tilde{\kappa}_1 n$ . In this case, Lemma 8.2 yields

$$(8.30) \quad -\log \text{VA}_1(gx, gx') \leq \tilde{\theta}_1$$

and Lemma 8.5 gives that

$$(8.31) \quad -\log d(gx, gx') \leq -\log d(x, x') - \log \text{VA}_1(x, x') - \hat{\kappa}_1 n.$$

Substituting (8.30) and (8.31) in (8.27) we obtain

$$\begin{aligned} -\log \text{VP}_1(gx, gx') &\leq \tilde{\theta}_1 - \gamma_1 \log d(x, x') - \gamma_1 \log \text{VA}_1(x, x') - \gamma_1 \hat{\kappa}_1 n \\ &\leq -\log \text{VP}_1(x, x') + \tilde{\theta}_1 + (1 - \gamma_1) \log \text{VA}_1(x, x') - \gamma_1 \hat{\kappa}_1 n. \end{aligned}$$

Since  $\gamma_1 \leq 1$ ,  $\text{VA}_1(x, x') \leq 1$ , and  $n \geq N_1 \geq 2\tilde{\theta}_1/(\gamma_1 \hat{\kappa}_1)$ , this yields

$$(8.32) \quad \begin{aligned} -\log \text{VP}_1(gx, gx') &\leq -\log \text{VP}_1(x, x') + \tilde{\theta}_1 - \gamma_1 \hat{\kappa}_1 n \\ &\leq -\log \text{VP}_1(x, x') - \frac{\gamma_1 \hat{\kappa}_1}{2} n \leq -\log \text{VP}_1(x, x') - \kappa'_1 n. \end{aligned}$$

The relations (8.29) and (8.32) contain the conclusion of Proposition 8.3.  $\square$

**8.3. The function  $-\log \psi_1$ .** Now we are going to prove that  $\psi_1 = \text{VP}_1$  satisfies Proposition 8.1. Let  $A = A(\nu_{\mathcal{O}})$  and  $B = B(\nu_{\mathcal{O}}) > 0$  be as in (8.2) and (8.3). Take  $\kappa'_1 > 0$ ,  $N_1 \in \mathbb{N}$  and  $\rho'_1 > 0$  as in Proposition 8.3. By definition,

$$(8.33) \quad -\log \psi_1(x, x') = -\log \text{VP}_1(x, x') = -\log \text{VA}_1(x, x') - \gamma_1 \log d(x, x')$$

for every  $x \neq x'$  in  $P$ . Define

$$(8.34) \quad C'_1 = B + \gamma_1 A.$$

Substituting (8.7) and (8.2) in (8.33), we find that

$$(8.35) \quad \begin{aligned} -\log \text{VP}_1(gx, gx') &\leq -\log \text{VA}_1(x, x') + Bn - \gamma_1 \log d(x, x') + \gamma_1 An \\ &\leq -\log \text{VP}_1(x, x') + C'_1 n \end{aligned}$$

for every  $x \neq x'$  in  $P$  and  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$ . Integrating (8.15) over  $\mathcal{E}'_1$  and (8.35) over the complement, and using the fact that  $\nu_{\mathcal{O}}^{(n)}((\mathcal{E}'_1)^c) < \delta$ , we get that

$$\int_G -\log \text{VP}_1(gx, gx') d\nu_{\mathcal{O}}^{(n)}(g) \leq -\log \text{VP}_1(x, x') - \kappa'_1 n + C'_1 \delta n.$$

for every  $n \geq N_1$  and  $x \neq x'$  in  $E_1(\rho'_1)$ . This completes the proof of Proposition 8.1.

Another relevant feature is that  $-\log \psi_1(x, x')$  goes to infinity when  $x$  and  $x'$  approach the equator  $E$ :

**Lemma 8.6.** *Given any  $R > 0$ , there exists  $\tilde{\rho}_1 = \tilde{\rho}_1(\nu_x, R) > 0$  such that  $-\log \psi_1(x, x') > R$  for any  $x \neq x'$  in  $E_1(\tilde{\rho}_1)$ .*

*Proof.* Consider any  $\tilde{\rho}_1 > 0$  and let  $v$  and  $v'$  be unit vectors in the direction of  $x$  and  $x'$ , respectively. Observe that

$$\frac{\|v^\perp\|}{\|v\|} = d(x, E) \leq \tilde{\rho}_1 \text{ and } \frac{\|(v')^\perp\|}{\|v'\|} = d(x', E) \leq \tilde{\rho}_1.$$

If  $d(x, x') \leq \tilde{\rho}_1^{1/2}$  then

$$(8.36) \quad \text{VP}_1(x, x') = \text{VA}_1(x, x')d(x, x')^{\gamma_1} \leq \tilde{\rho}_1^{\gamma_1/2}.$$

Now suppose that  $d(x, x') = |\sin \angle(v, v')|$  is greater than  $\tilde{\rho}_1^{1/2}$ . Then every  $u$  in the great circle  $y$  generated by  $x$  and  $x'$  may be written as  $u = \alpha v + \alpha' v'$  with  $\alpha, \alpha' \in \mathbb{R}$  such that  $\|u\| \geq \tilde{\rho}_1^{1/2} \|\alpha v\|$  and  $\|u\| \geq \tilde{\rho}_1^{1/2} \|\alpha' v'\|$ . Thus,

$$\frac{\|u^\perp\|}{\|u\|} \leq \frac{\|\alpha v^\perp\|}{\|u\|} + \frac{\|\alpha' (v')^\perp\|}{\|u\|} \leq \frac{1}{\tilde{\rho}_1^{1/2}} \left( \frac{\|v^\perp\|}{\|v\|} + \frac{\|(v')^\perp\|}{\|v'\|} \right) \leq 2\tilde{\rho}_1^{1/2}.$$

Since  $u$  is arbitrary, this proves that  $\text{VA}_1(x, x') \leq 2\tilde{\rho}_1^{1/2}$ , and so

$$(8.37) \quad \text{VP}_1(x, x') = \text{VA}_1(x, x')d(x, x')^{\gamma_1} \leq 2\tilde{\rho}_1^{1/2} \leq 2\tilde{\rho}_1^{\gamma_1/2}$$

(because  $\gamma_1 \leq 1$ ). The inequalities (8.36) and (8.37) imply that

$$-\log \psi_1(x, x') = -\log \text{VP}_1(x, x') \geq -\frac{\gamma_1}{2} \log \tilde{\rho}_1 - \log 2$$

for any  $x \neq x'$  in  $E_1(\tilde{\rho}_1)$ , which yields the claim.  $\square$

## 9. STABILIZATION AND CUT-OFF

Previously, we constructed a function  $-\log \psi_1 = -\log \text{VP}_1$  that exhibits some of the features of a Margulis function. In this section we modify this function to correct two important defects.

One problem with  $-\log \psi_1$ , that originates from  $\text{VA}_1$ , is that the definition refers explicitly to  $E$ . The reason why this is a problem is that for  $k \in \mathbb{N}$  the equator is not necessarily a  $\nu_k$ -invariant set. Thus, even for  $k$  large, for  $g \in \text{supp } \nu_k$  there is no uniform upper bound on the absolute value of

$$-\log \psi_1(gx, gx') + \log \psi_1(x, x')$$

because  $-\log \text{VA}_1(gx, gx')$  may be wildly different from  $-\log \text{VA}_1(x, x')$ . To rectify this, in Section 9.1 we define “stabilized” versions of the functions  $\text{VA}_1$  and  $\text{VP}_1$ . This will come at a price: the analogue of Proposition 8.1 will hold only outside a “stabilization region” near  $E$ . Still, as we will explain in Section 10, stabilization does allow us to by-pass this first difficulty in a satisfactory way.

Another problem is that, because of the way the Markov operators  $\mathcal{T}_{k,1}$  will be constructed, we have little control over the border region. The simplest way to by-pass this is to take the Margulis function to be bounded when either  $x$  or  $x'$  is in the border region. The function  $-\log \psi_1$  does not satisfy this: for instance, it can get arbitrarily large when  $x$  and  $x'$  are close to each other and, of course, that may occur even if they are both in the border region. Thus, it is necessary to cut that function off in (a neighborhood of) the border region.

However, doing a cut-off creates a discontinuity that translates into a drastic failure of (7.14) near the discontinuity. The main idea to handle this, which we

call *recoupling*, will be explained in Section 11.1. The recoupling technique is far from universal, it can only handle certain types of discontinuities. Thus, it makes sense to make the discontinuity as mild as possible. In Section 9.2 we explain just how to do this. In particular, we only cut  $-\log \psi_1$  off if  $x$  and  $x'$  are both in (a neighborhood of) the border region. In the present ( $r = 1$ ) situation this is fine because if one of the points,  $x$  or  $x'$ , is close to the border region and the other one is far outside then  $-\log \psi_1(x, x')$  is bounded anyway. The steps  $r > 1$  involve additional issues, which we will discuss in Sections 13 and 15.

We use the following elementary inequalities, whose proof we leave to the reader: given any  $C > 0$ ,  $c > 0$  and  $\psi > 0$ ,

$$(9.1) \quad \log(\Omega + c\psi^{-1}) \leq \log(\Omega + \psi^{-1}) + \log c \quad \text{if } c > 1$$

$$(9.2) \quad \log(\Omega + c\psi^{-1}) \leq \log(\Omega + \psi^{-1}) \quad \text{if } c < 1$$

$$(9.3) \quad \log(\Omega + c\psi^{-1}) \leq \log(\Omega + \psi^{-1}) + \log \sqrt{c} \quad \text{if } c < 1 \text{ and } \psi^{-1} \geq \Omega/\sqrt{c}.$$

Let  $\kappa'_1 = \kappa'_1(\nu_x) > 0$ ,  $C'_1 = C'_1(\nu_x) > 0$ ,  $N_1 = N_1(\nu_x, \delta) \in \mathbb{N}$  be as in Propositions 8.1 and 8.3. Keep in mind that  $\varepsilon_1 < \rho'_1 \leq \hat{\rho}_1$  and  $C'_1 = B + \gamma_1 A$ , according to (8.26), (8.34) and (9.12).

**9.1. Stabilization.** Let  $B = B(\nu_x) > 0$  and  $\gamma_1 = \gamma_1(\nu_x) > 0$  be as in (8.3) and (8.26), respectively. For each  $\omega > 0$ , define the *stabilized vertical angle*

$$(9.4) \quad \text{SVA}_1(x, x'; \omega) = \max \{ \text{VA}_1(x, x'), \omega e^{-Bn} \}$$

and the *stabilized vertical projection* by

$$(9.5) \quad \psi_1(x, x'; \omega) = \text{SVP}_1(x, x'; \omega) = \text{SVA}_1(x, x'; \omega) d(x, x')^{\gamma_1}$$

for every  $x \neq x'$  in  $E_1(\varepsilon_1)$ . The estimate (8.35) remains valid for the stabilized vertical projection:

**Lemma 9.1.** *For every  $x \neq x'$  in  $P$ ,  $g \in \text{supp } \nu_x^{(n)}$ , and  $\omega > 0$ ,*

$$(9.6) \quad -\log \psi_1(gx, gx'; \omega) \leq -\log \psi_1(x, x'; \omega) + C'_1 n.$$

*Proof.* We begin by claiming that

$$(9.7) \quad -\log \text{SVA}_1(gx, gx'; \omega) \leq -\log \text{SVA}_1(x, x'; \omega) + Bn.$$

The proof can be split into two cases. First, suppose that  $\text{VA}_1(x, x') < \omega$ . Then, by the definition (9.4),

$$\begin{aligned} -\log \text{SVA}_1(x, x'; \omega) &\geq -\log \omega \text{ and} \\ -\log \text{SVA}_1(gx, gx'; \omega) &\leq -\log \omega + Bn \leq -\log \text{SVA}_1(x, x'; \omega) + Bn, \end{aligned}$$

as claimed. Now suppose that  $\text{VA}_1(x, x') \geq \omega$ . Then, again by (9.4),

$$\begin{aligned} -\log \text{SVA}_1(x, x'; \omega) &= -\log \text{VA}_1(x, x') \text{ and} \\ -\log \text{SVA}_1(gx, gx'; \omega) &\leq -\log \text{VA}_1(gx, gx'). \end{aligned}$$

Then (9.7) is a direct consequence of (8.7). This completes the proof of (9.7). Now (9.6) follows easily from (9.7) and (8.2): recalling the definition of  $C'_1$  in (8.34),

$$(9.8) \quad \begin{aligned} -\log \text{SVP}_1(gx, gx'; \omega) &= -\log \text{SVA}_1(gx, gx'; \omega) - \gamma_1 \log d(gx, gx') \\ &\leq -\log \text{SVA}_1(x, x'; \omega) + Bn - \gamma_1 \log d(x, x') - \gamma_1 An \\ &= -\log \text{SVP}_1(x, x') + (B + \gamma_1 A)n. \end{aligned}$$

This completes the argument.  $\square$

The *stabilization region* is the set of pairs  $(x, x')$  such that  $\text{VA}_1(x, x') < \omega$ . We have seen in (8.7) that

$$\text{VA}_1(gx, gx') \geq \text{VA}_1(x, x')e^{-Bn}$$

for any  $x \neq x'$  in  $P$ ,  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$  and  $n \in \mathbb{N}$ . Consequently, if  $(x, x')$  is not in the stabilization region then

$$(9.9) \quad \begin{aligned} \text{SVA}_1(x, x'; \omega) &= \text{VA}_1(x, x') \text{ and } \text{SVA}_1(gx, gx'; \omega) = \text{VA}_1(gx, gx') \\ \text{SVP}_1(x, x'; \omega) &= \text{VP}_1(x, x') \text{ and } \text{SVP}_1(gx, gx'; \omega) = \text{VP}_1(gx, gx') \end{aligned}$$

for any  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$  and  $n \in \mathbb{N}$ .

Propositions 8.1 and 8.3 immediately yield the following analogues for stabilized vertical angles and vertical projections:

**Proposition 9.2.** *For every  $\delta > 0$ ,  $n \geq N_1$ ,  $x \neq x'$  in  $E_1(\rho'_1)$ , and  $\omega > 0$  with  $\text{VA}_1(x, x') \geq \omega$ ,*

$$(9.10) \quad -\log \psi_1(gx, gx'; \omega) \leq -\log \psi_1(x, x'; \omega) - \kappa'_1 n \text{ for every } g \in \mathcal{E}'_1,$$

and

$$(9.11) \quad \int_G -\log \psi_1(gx, gx'; \omega) d\nu_{\mathcal{O}}^{(n)}(g) \leq -\log \psi_1(x, x'; \omega) - (\kappa'_1 - C'_1 \delta)n.$$

*Proof.* Since it is assumed that  $(x, x')$  is not in the stabilization region, (9.10) is a restatement of Proposition 8.3 and (9.11) is a restatement of Proposition 8.1.  $\square$

**9.2. Cutoff.** Recall that  $\varepsilon_1 = \varepsilon_1(\nu_{\mathcal{O}}, \delta, n) > 0$  was chosen in Section 7.2, in the context of (7.3). As observed then, it may be taken to be arbitrarily small. In particular, it is no restriction to suppose that

$$(9.12) \quad \varepsilon_1 < \min\{\rho_0, \rho'_1\}$$

where  $\rho_0 = \rho_0(\nu_{\mathcal{O}}, n) > 0$  is as in Corollaries 5.6 and 5.7 and  $\rho'_1 = \rho'_1(\nu_{\mathcal{O}}, \delta, n) > 0$  is as in Propositions 8.1 and 8.3.

Arguing twice as in Remark 7.3, we find constants  $\varepsilon'_1 = \varepsilon'_1(\nu_{\mathcal{O}}, \delta, n) > 0$  and  $\tilde{\varepsilon}_1 = \tilde{\varepsilon}_1(\nu_{\mathcal{O}}, \delta, n) > 0$  with  $0 < \varepsilon'_1 < \tilde{\varepsilon}_1 < \varepsilon_1$ , and a compact neighborhood  $\mathcal{W}_1 = \mathcal{W}_1(\nu_{\mathcal{O}}, \delta, n)$  of  $\text{supp } \nu_{\mathcal{O}}^{(n)}$ , such that

$$(9.13) \quad gx \in E_1(\varepsilon_1/2) \text{ for every } x \in E_1(2\tilde{\varepsilon}_1) \text{ and } g \in \mathcal{W}_1 \text{ and}$$

$$(9.14) \quad g^{-1}x \in E_1(\tilde{\varepsilon}_1/2) \text{ for every } x \in E_1(2\varepsilon'_1) \text{ and } g \in \mathcal{W}_1.$$

Let  $\kappa_0 = \kappa_0(\nu_{\mathcal{O}}) > 0$  be as in Proposition 5.5, and define  $\varepsilon''_1 = \varepsilon''_1(\nu_{\mathcal{O}}, \delta, n)$  by

$$(9.15) \quad \varepsilon''_1 = 3\varepsilon'_1 e^{-\kappa_0 n/2}.$$

Taking  $\rho = \varepsilon''_1$  in Corollary 5.7, and recalling that  $\varepsilon_1$  was chosen smaller than  $\rho_0$ , we get that there are  $\tilde{k}_1 = \tilde{k}_1(\nu_{\mathcal{O}}, \delta, n) \in \mathbb{N}$  and  $\mathcal{D}_k(x) = \mathcal{D}_k(\nu_{\mathcal{O}}, \delta, n, x) \subset \text{supp } \nu_k^{(n)}$  such that  $\nu_k^{(n)}(\mathcal{D}_k(x)^c) < \delta$  and

$$d(gx, E) > e^{\kappa_0 n/2} d(x, E) > e^{\kappa_0 n/2} \varepsilon''_1 > 2\varepsilon'_1$$

for any  $g \in \mathcal{D}_k(x)$ ,  $x \in E(\varepsilon_1, \varepsilon''_1)$  and  $k \geq \tilde{k}_1$ . In other words, for  $k \geq \tilde{k}_1$ ,

$$(9.16) \quad x \in E(\varepsilon_1, \varepsilon''_1) \Rightarrow gx \notin E(2\varepsilon'_1) \text{ for every } g \in \mathcal{D}_k(x).$$

Increasing  $\tilde{k}_1$  if necessary, we may suppose that  $\text{supp } \nu_k^{(n)} \subset \mathcal{W}_1$  for every  $k \geq \tilde{k}_1$ . Then (9.13) and (9.14) imply

$$(9.17) \quad E_1(2\tilde{\varepsilon}_1) \subset \mathcal{X}_{\nu_k^{(n)}} E_1(\varepsilon_1) \text{ and } E_1(2\varepsilon'_1) \subset \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_1)$$

$$(9.18) \quad x \notin E_1(\tilde{\varepsilon}_1/2) \Rightarrow gx \notin E_1(2\varepsilon'_1) \text{ for every } g \in \text{supp } \nu_k^{(n)}.$$

Finally, fix  $\Omega_1 = \Omega_1(\nu_\infty, \delta, n) > 1$  large enough that

$$(9.19) \quad \Omega_1 \varepsilon'_1 (\varepsilon'_1 - 2\varepsilon''_1)^{\gamma_1} \geq 1$$

and define

$$(9.20) \quad \Psi_1(x, x'; \omega) = \begin{cases} \log(\Omega_1 + \psi_1(x, x'; \omega)^{-1}) & \text{if } x \in E_1(2\varepsilon''_1) \text{ or } x' \in E_1(2\varepsilon''_1) \\ \log \Omega_1 & \text{otherwise.} \end{cases}$$

We will refer to the set  $E_1(2\varepsilon''_1)^c \times E_1(2\varepsilon''_1)^c$  as the *cut-off region*. See Figure 7.

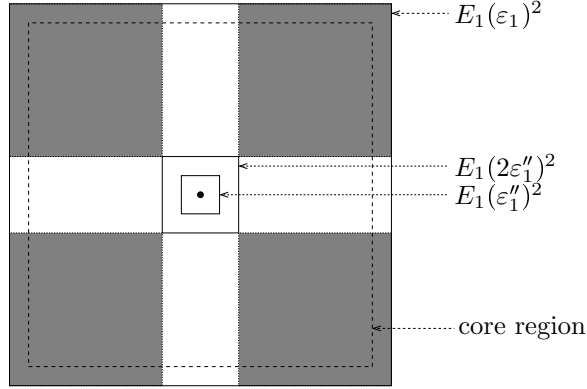


FIGURE 7. Illustrating the cut-off in the definition of the Margulis function. The black dot at the center marks the point  $(E, E)$ . The dashed line represents the boundary between the  $\nu_k^{(n)}$ -core and the  $\nu_k^{(n)}$ -border of  $E_1(\varepsilon_1)$ . The shaded area is the cut-off region, where  $\Psi_1 \equiv \log \Omega_1$ .

**Proposition 9.3.** *There exist  $\kappa_1''' = \kappa_1'''(\nu_\infty) > 0$  and for each  $\delta > 0$  and  $n \geq N_1$  there exists  $\varepsilon_1''' = \varepsilon_1'''(\nu_\infty, \delta, n) > 0$  such that*

(i) *For any  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $\Psi_1(x, x'; \omega) > \log \Omega_1$ ,*

$$\int_G \Psi_1(gx, gx'; \omega) d\nu_\infty^{(n)}(g) \leq \Psi_1(x, x'; \omega) + C'_1 n.$$

(ii) *For any  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $\Psi_1(x, x'; \omega) > \log \Omega_1$  and  $\text{VA}_1(x, x') \geq \omega$ ,*

$$\int_G \Psi_1(gx, gx'; \omega) d\nu_\infty^{(n)}(g) \leq \Psi_1(x, x'; \omega) + C'_1 \delta n.$$

(iii) *For any  $x \neq x'$  in  $E_1(\varepsilon_1''')$  with  $\text{VA}_1(x, x') \geq \omega$ ,*

$$\int_G \Psi_1(gx, gx'; \omega) d\nu_\infty^{(n)}(g) \leq \Psi_1(x, x'; \omega) - (\kappa_1''' - C'_1 \delta) n.$$

*Proof.* Define  $\kappa_1''' = \kappa_1'/2$ . Let  $n \geq N_1$ . Part (i) of the proposition is a consequence of the following lemma:

**Lemma 9.4.** *If  $x \neq x'$  in  $P$  are such that  $\Psi_1(x, x'; \omega) > \log \Omega_1$  then*

$$(9.21) \quad \Psi_1(gx, gx'; \omega) \leq \Psi_1(x, x'; \omega) + C_1' n \text{ for any } g \in \text{supp } \nu_\omega^{(n)}.$$

*Proof.* By (9.6) and (9.1),

$$\begin{aligned} \Psi_1(gx, gx'; \omega) &\leq \log(\Omega_1 + \psi_1(gx, gx'; \omega)^{-1}) \leq \log(\Omega_1 + e^{C_1' n} \psi_1(x, x'; \omega)^{-1}) \\ &\leq \log(\Omega_1 + \psi_1(x, x'; \omega)^{-1}) + C_1' n = \Psi_1(x, x'; \omega) + C_1' n. \end{aligned}$$

This gives the claim.  $\square$

To prove part (ii) we use Proposition 9.2: given any  $\omega > 0$  and  $x \neq x'$  in  $E_1(\varepsilon_1)$ ,

$$(9.22) \quad -\log \psi_1(gx, gx'; \omega) \leq -\log \psi_1(x, x'; \omega) - \kappa_1' n \leq -\log \psi_1(x, x'; \omega)$$

for every  $g \in \mathcal{E}'_1$ . Then,

$$(9.23) \quad \begin{aligned} \Psi_1(gx, gx'; \omega) &\leq \log(\Omega_1 + \psi_1(gx, gx'; \omega)^{-1}) \\ &\leq \log(\Omega_1 + \psi_1(x, x'; \omega)^{-1}) = \Psi_1(x, x'; \omega), \end{aligned}$$

for every  $g \in \mathcal{E}'_1$ . Integrating (9.23) over  $\mathcal{E}'_1$  and (9.21) over the complement, we obtain the estimate in part (ii).

Now take  $c = e^{-2\kappa_1''' n}$  in the relation (9.3). By Lemma 8.6, there exists  $\varepsilon_1''' > 0$  depending only on  $\nu_x$ ,  $\delta$  and  $n$  (through  $c$  and  $\Omega_1$ ) such that

$$(9.24) \quad -\log \psi_1(x, x'; \omega) = -\log \psi_1(x, x') \geq \log(\Omega_1/\sqrt{c})$$

for any  $x \neq x'$  in  $E_1(\varepsilon_1''')$ . Then, using (9.22), (9.24), and (9.3),

$$(9.25) \quad \begin{aligned} \Psi_1(gx, gx'; \omega) &\leq \log(\Omega_1 + \psi_1(gx, gx'; \omega)^{-1}) \\ &\leq \log(\Omega_1 + e^{-2\kappa_1''' n} \psi_1(x, x'; \omega)^{-1}) \\ &\leq \log(\Omega_1 + \psi_1(x, x'; \omega)^{-1}) - \kappa_1''' n = \Psi_1(x, x'; \omega) - \kappa_1''' n \end{aligned}$$

for every  $g \in \mathcal{E}'_1$ . Integrating (9.25) over  $\mathcal{E}'_1$  and (9.21) over the complement, we obtain the estimate in part (iii) of the proposition.  $\square$

**Lemma 9.5.** *For any  $\omega > 0$  and  $x \neq x'$  in  $E_1(\varepsilon_1)$  such that  $x \notin E_1(\varepsilon'_1)$  or  $x' \notin E_1(\varepsilon'_1)$ ,*

$$\log \Omega_1 \leq \Psi_1(x, x'; \omega) \leq \log \Omega_1 + \log 2.$$

*Proof.* The inequality on the left is an immediate consequence of the definition (9.20), and the same is true for the one on the right if both points  $x$  and  $x'$  are outside  $E_1(2\varepsilon_1'')$ . Let us suppose that  $x \notin E_1(\varepsilon'_1)$  but  $x' \in E_1(2\varepsilon_1'')$ ; the case when  $x \in E_1(2\varepsilon_1'')$  but  $x' \notin E_1(\varepsilon'_1)$  is analogous. Let  $y = y(x, x')$  be the great circle associated to  $x$  and  $x'$ . Then

$$d(x, x') \geq \varepsilon'_1 - 2\varepsilon_1'' \text{ and } \text{VA}_1(x, x') = d(y, E) \geq d(x, E) > \varepsilon'_1$$

and so

$$\psi_1(x, x'; \omega) \geq \psi_1(x, x') \geq \varepsilon'_1(\varepsilon'_1 - 2\varepsilon_1'')^{\gamma_1}.$$

In view of the definition of  $\Omega_1$  in (9.19), it follows that

$$\Psi_1(x, x'; \omega) = \log(\Omega_1 + \psi_1(x, x'; \omega)^{-1}) \leq \log 2\Omega_1,$$

as claimed.  $\square$



## 10. TURNING THE PERTURBATION ON

Now we show that the conclusions of Proposition 9.3 hold for  $\nu_k^{(n)}$  instead of  $\nu_\infty^{(n)}$ , as long as  $k \in \mathbb{N}$  is sufficiently large. More precisely, we prove:

**Proposition 10.1.** *Given  $\delta > 0$ ,  $n \geq N_1$ , and  $\omega > 0$  there is  $k_1 = k_1(\nu_\infty, \delta, n, \omega) \in \mathbb{N}$  such that the following holds for every  $k \geq k_1$ :*

(i) *For any  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $\Psi_1(x, x'; \omega) > \log \Omega_1$ ,*

$$\int_G \Psi_1(gx, gx'; \omega) d\nu_k^{(n)}(g) \leq \Psi_1(x, x'; \omega) + C'_1 n.$$

(ii) *For any  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $\Psi_1(x, x'; \omega) > \log \Omega_1$  and  $\text{VA}_1(x, x') \geq \omega$ ,*

$$\int_G \Psi_1(gx, gx'; \omega) d\nu_k^{(n)}(g) \leq \Psi_1(x, x'; \omega) + C'_1 \delta n.$$

(iii) *For any  $x \neq x'$  in  $E_1(\varepsilon_1''')$  with  $\text{VA}_1(x, x') \geq \omega$ ,*

$$\int_G \Psi_1(gx, gx'; \omega) d\nu_k^{(n)}(g) \leq \Psi_1(x, x'; n) - (\kappa'_1 - C'_1 \delta) n.$$

Keep in mind that we have chosen  $\varepsilon_1 < \rho'_1 \leq \hat{\rho}_1$  and  $C'_1 = B + \gamma_1 A$ . Recall also that  $\tilde{k}_1 = \tilde{k}_1(\nu_\infty, \delta, n) \in \mathbb{N}$  was chosen so that the relations (9.16) through (9.18) hold for every  $k \geq \tilde{k}_1$ . Moreover,  $N_1 \in \mathbb{N}$  is given by Propositions 8.3 and 9.2.

*Proof.* We are going to extend to large  $k \in \mathbb{N}$  several estimates in the proof of Proposition 9.3. This will require a number of conditions on  $k$ , depending on  $\nu_\infty$ ,  $\delta$ ,  $n$  and  $\omega$ , that we state along the way. We begin with the following extension of Lemma 9.4:

**Lemma 10.2.** *Given  $n \geq N_1$  and  $\omega > 0$ , there is  $\hat{k}_1 = \hat{k}_1(\nu_\infty, n, \omega) \in \mathbb{N}$  such that if  $x \neq x'$  in  $P$  are such that  $\Psi_1(x, x'; \omega) > \log \Omega_1$  then*

$$(10.1) \quad \Psi_1(gx, gx'; \omega) \leq \Psi_1(x, x'; \omega) + C'_1 n$$

for any  $g \in \text{supp } \nu_k^{(n)}$  and any  $k \geq \hat{k}_1$ .

*Proof.* Let  $n \geq N_1$  and  $\omega > 0$ . We claim that there is  $\hat{k}_1 = \hat{k}_1(\nu_\infty, n, \omega)$  such that

$$(10.2) \quad -\log \text{SVA}_1(gx, gx'; \omega) \leq -\log \text{SVA}_1(x, x'; \omega) + Bn$$

for any  $x \neq x'$ ,  $g \in \text{supp } \nu_k^{(n)}$ , and  $k \geq \hat{k}_1$ . This can be seen as follows. If  $\text{VA}_1(x, x') < \omega$  then, by the definition (9.4),

$$\begin{aligned} -\log \text{SVA}_1(x, x'; \omega) &\geq -\log \omega \text{ and} \\ -\log \text{SVA}_1(gx, gx'; \omega) &\leq -\log \omega + Bn \leq -\log \text{SVA}_1(x, x'; \omega) + Bn, \end{aligned}$$

as claimed. Now suppose that  $\text{VA}_1(x, x') \geq \omega$ . The relation (8.7) does not apply here. Instead, from (8.5) and (9.4) we get that

$$(10.3) \quad -\log \text{SVA}_1(gx, gx'; \omega) \leq -\log \text{SVA}_1(x, x'; \omega) - \log \frac{\|(gu)^\perp\| \|u\|}{\|gu\| \|u^\perp\|}$$

for every  $g \in \text{supp } \nu_k^{(n)}$ , with  $u = u(x, x')$  realizing the supremum in (8.4). The assumption  $\text{VA}_1(x, x') \geq \omega$  means that  $\|u^\perp\| \geq \omega \|u\|$ . Since  $\text{supp } \nu_k$  converges to

supp  $\nu_\infty$  in the Hausdorff topology, we may find  $r_k = r_k(\nu_\infty, n) \rightarrow 0$  such that every  $g \in \text{supp } \nu_k^{(n)}$  is in the  $r_k$ -neighborhood of some  $f \in \text{supp } \nu_\infty^{(n)}$ . Then

$$\|(gu)^\perp - (fu)^\perp\| \leq \|gu - fu\| \leq r_k \|u\| \leq \frac{r_k}{\omega} \|u^\perp\|.$$

Fix  $\hat{k}_1 \in \mathbb{N}$  large enough that, for every  $g \in \text{supp } \nu_k^{(n)}$  and  $k \geq \hat{k}_1$ ,

$$\begin{aligned} -\log \frac{\|(gu)^\perp\| \|u\|}{\|gu\| \|u^\perp\|} &\leq -\log \frac{\|(fu)^\perp\| \|u\|}{\|fu\| \|u^\perp\|} + \log 2 \\ &= -\log \frac{\|f^\perp u^\perp\| \|u\|}{\|fu\| \|u^\perp\|} + \log 2 \\ &\leq \log \|(f^\perp)^{-1}\| + \log \|f\| + \log 2 \leq Bn. \end{aligned}$$

Substituting this in (10.3) completes the proof of (10.2).

Next, substituting (10.2) and (8.2) in the definition (9.5) we find that

$$(10.4) \quad \begin{aligned} -\log \psi_1(gx, gx'; \omega) &\leq -\log \text{SVA}_1(x, x'; \omega) + Bn - \log d(x, x') + \gamma_1 An \\ &= -\log \psi_1(x, x'; \omega) + C'_1 n \end{aligned}$$

for any  $x \neq x'$ ,  $g \in \text{supp } \nu_k^{(n)}$ , and  $k \geq \hat{k}_1$ . Using (9.1), it follows that

$$\begin{aligned} \Psi_1(gx, gx'; \omega) &\leq \log (\Omega_1 + \psi_1(gx, gx'; \omega)^{-1}) \\ &\leq \log (\Omega_1 + e^{C'_1 n} \psi_1(x, x'; \omega)^{-1}) \\ &\leq \log (\Omega_1 + \psi_1(x, x'; \omega)^{-1}) + C'_1 n = \Psi_1(x, x'; \omega) + C'_1 n, \end{aligned}$$

as stated.  $\square$

Next, we prove the following extension of Lemma 8.2:

**Lemma 10.3.** *Given  $\delta > 0$ ,  $n \geq N_1$  and  $\omega > 0$ , there is  $k'_1 = k'_1(\nu_\infty, \delta, n, \omega) \in \mathbb{N}$  and for any  $x \neq x'$  in  $P$  with  $\text{VA}_1(x, x') \geq \omega$  there is  $\mathcal{E}'_{k,1} = \mathcal{E}'_{k,1}(\nu_\infty, \delta, n, x, x', \omega) \subset \text{supp } \nu_k^{(n)}$  with  $\nu_k^{(n)}((\mathcal{E}'_{k,1})^c) < \delta$  and*

$$(10.5) \quad -\log \text{SVA}_1(gx, gx'; \omega) \leq \max\{-\log \text{SVA}_1(x, x'; \omega) - \tilde{\kappa}_1 n, \tilde{\theta}_1\}$$

for every  $g \in \mathcal{E}'_{k,1}$  and  $k \geq k'_1$ .

*Proof.* Fix  $\delta > 0$  and  $n \geq N_1$  and  $\omega > 0$ . Let  $P_\omega$  denote the (compact) subset of all  $v \in P$  such that  $\|v^\perp\|/\|v\| \geq \omega/2$ . For  $v \in P_\omega$  and  $g$  in some compact neighborhood  $V_\omega$  of  $\text{supp } \nu_\infty^{(n)}$ , consider

$$(10.6) \quad (v, g) \mapsto -\log \frac{\|(gv)^\perp\|}{\|gv\|}.$$

As long as  $V_\omega$  is sufficiently small, depending on  $\nu_\infty$ ,  $n$  and  $\omega$ , the map (10.6) is well defined and (uniformly) continuous. So, there exists  $\alpha = \alpha(\nu_\infty, n, \omega) > 0$  such that

$$(10.7) \quad -\log \frac{\|(gu)^\perp\|}{\|gu\|} \leq -\log \frac{\|(fv)^\perp\|}{\|fv\|} + \log 2$$

whenever  $d(u, v) < \alpha$  and  $d(g, f) < \alpha$ . Reducing  $\alpha$  if necessary, depending only on  $\omega$ , we may also assume that

$$(10.8) \quad d(u, v) < \alpha \quad \Rightarrow \quad -\log \frac{\|u^\perp\|}{\|u\|} \geq -\log \frac{\|v^\perp\|}{\|v\|} - \log 2.$$

Fix  $v_1, \dots, v_l \in P_\omega$  such that  $P_\omega \subset B(v_1, \alpha) \cup \dots \cup B(v_l, \alpha)$ . For each  $v \in P_\omega$  choose  $j \in \{1, \dots, l\}$  such that  $v \in B(v_j, \alpha)$  and then define  $\mathcal{E}_{k,0} = \mathcal{E}_{k,0}(\nu_\infty, \delta, n, v, \omega) \subset \text{supp } \nu_k^{(n)}$  by

$$(10.9) \quad \mathcal{E}_{k,0} = [\alpha\text{-neighborhood of } \mathcal{E}_0(\nu_\infty, \delta, n, v_j^\perp)] \cap \text{supp } \nu_k^{(n)},$$

where  $\mathcal{E}_0(\nu_\infty, \delta, n, v_j^\perp) \subset \text{supp } \nu_\infty^{(n)}$  is as defined in Proposition 5.5. Since  $\nu_k^{(n)}$  converges to  $\nu_\infty^{(n)}$  in the weak\* topology, the limit inferior of the  $\nu_k^{(n)}$ -measure of (10.9) as  $k \rightarrow \infty$  is greater than or equal to

$$\nu_\infty^{(n)}(\mathcal{E}_0(\nu_\infty, \delta, n, v_j^\perp)) > 1 - \delta$$

for every  $j = 1, \dots, l$ . In particular, there is  $k'_1 = k'_1(\nu_\infty, \delta, n, \omega) \in \mathbb{N}$  such that

$$(10.10) \quad \nu_k^{(n)}(\mathcal{E}_{k,0}) > 1 - \delta \text{ for every } k \geq k'_1 \text{ and } v \in P_\omega.$$

Given  $x \neq x'$  with  $\text{VA}_1(x, x') \geq \omega$ , take  $u = u^E + u^\perp$  to be a non-zero vector that realizes the supremum in (8.4). Then,

$$(10.11) \quad \omega \leq \text{VA}_1(x, x') = \frac{\|u^\perp\|}{\|u\|} \quad (\text{in particular, } u \in P_\omega) \text{ and}$$

$$(10.12) \quad \text{VA}_1(gx, gx') \geq \frac{\|(gu)^\perp\|}{\|gu\|} \geq \text{VA}_1(x, x') \frac{\|(gu)^\perp\|}{\|gu\|} \frac{\|u\|}{\|u^\perp\|}$$

for any  $g \in G$ . Then define

$$(10.13) \quad \mathcal{E}'_{k,1} = \mathcal{E}_{k,0}(\nu_\infty, \delta, n, u, \omega).$$

It follows from (10.10) that  $\nu_k^{(n)}((\mathcal{E}'_{k,1})^c) < \delta$  for every  $k \geq k'_1$ .

Let  $g \in \mathcal{E}'_{k,1}$  and  $k \geq k'_1$ . Then, by definition, there exist  $v = v^E + v^\perp$  in  $P_\omega$  (take  $v = v_j$  as in (10.9)) and  $f \in \mathcal{E}_0(\nu_\infty, \delta, n, v^\perp) \subset \text{supp } \nu_\infty^{(n)}$  such that  $d(u, v) < \alpha$  and  $d(g, f) < \alpha$ . Thus, substituting (10.7) and (10.8) in (10.12), we find that

$$(10.14) \quad \text{VA}_1(gx, gx') \geq \frac{1}{2} \frac{\|f^\perp v^\perp\|}{\|fv\|} \geq \frac{1}{4} \text{VA}_1(x, x') \frac{\|f^\perp v^\perp\|}{\|fv\|} \frac{\|v\|}{\|v^\perp\|}.$$

Let  $\tau_0 = \tau_0(\nu_\infty, \delta) > 0$  be as in Proposition 5.5. If  $\|f^\perp v^\perp\|/\|fv\| \geq \tau_0/2$  then the first part of (10.14) gives that (recall (8.9) also)

$$(10.15) \quad -\log \text{VA}_1(gx, gx') \leq -\log \frac{\tau_0}{4} \leq \tilde{\theta}_1.$$

Now suppose that  $\|f^\perp v^\perp\|/\|fv\| < \tau_0/2$ . Then part (2) of Proposition 5.5 gives that

$$(10.16) \quad \frac{\|fv^\perp\|}{\|fv\|} < \frac{1}{2} \text{ and so } \frac{\|fv^E\|}{\|fv\|} > \frac{1}{2}.$$

Substituting (10.16) and  $\|v\| \geq \|v^E\|$  in (10.14), we find that

$$(10.17) \quad \begin{aligned} \text{VA}_1(gx, gx') &\geq \frac{1}{8} \text{VA}_1(x, x') \frac{\|f^\perp v^\perp\|}{\|v^\perp\|} \frac{\|v^E\|}{\|fv^E\|} \\ &= \frac{1}{8} \text{VA}_1(x, x') \frac{\|Df_{v^E}^\perp v^\perp\|}{\|v^\perp\|}. \end{aligned}$$

By part (1) of Proposition 5.5 and the choices of  $\tilde{\kappa}_1 > 0$  and  $\tilde{N}_1 \in \mathbb{N}$  in (8.9), this implies

$$(10.18) \quad \begin{aligned} -\log \text{VA}_1(gx, gx') &\leq -\log \text{VA}_1(x, x') + \log 8 - 2\tilde{\kappa}_1 n \\ &\leq -\log \text{VA}_1(x, x') - \tilde{\kappa}_1 n. \end{aligned}$$

The conclusion of the lemma is contained in (10.15) and (10.18).  $\square$

Next, let us prove the following extension of Lemma 8.5:

**Lemma 10.4.** *Given  $\delta > 0$  and  $n \geq N_1$  there exists  $k_1'' = k_1''(\nu_\infty, \delta, n) \in \mathbb{N}$  such that for any  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $-\log \text{VA}_1(x, x') \leq \tilde{\theta}_1 + \tilde{\kappa}_1 n$  there exists  $\mathcal{E}_{k,1}'' = \mathcal{E}_{k,1}''(\nu_\infty, \delta, n, x, x') \subset \text{supp } \nu_k^{(n)}$  with  $\nu_k^{(n)}((\mathcal{E}_{k,1}'')^c) < \delta$  and*

$$(10.19) \quad -\log d(gx, gx') \leq -\log d(x, x') - \log \text{VA}_1(x, x') - \hat{\kappa}_1 n$$

for every  $g \in \mathcal{E}_{k,1}''$  and  $k \geq k_1''$ .

*Proof.* Let  $\hat{P}$  denote the (compact) subset of pairs  $(v, w) \in P \times P$  such that

$$(10.20) \quad \frac{\|v^\perp\|}{\|v\|} \leq \hat{\rho}_1 < 2\hat{\rho}_1 \leq \frac{\|w^\perp\|}{\|w\|}$$

(as before,  $v^\perp$  and  $w^\perp$  denote the vertical components of  $v$  and  $w$ ). Let  $n \geq N_1$  and  $\hat{V}$  be a compact neighborhood of the support of  $\nu_\infty^{(n)}$ . Since,  $\hat{\rho}_1 = \hat{\rho}_1(\nu_\infty, n)$ ,  $\hat{P}$  and  $\hat{V}$  depend only on  $\nu_\infty$  and  $n$ . Condition (10.20) ensures that the angle between  $v$  and  $w$  is bounded away from zero and, consequently, so is the angle between  $gv$  and  $gw$  for any  $g \in \hat{V}$ ; both bounds depend only on  $\nu_\infty$  and  $n$ . Thus, the map

$$(10.21) \quad (v, w, g) \mapsto -\log \frac{\|\Pi_{gv} gw\|}{\|gv\|} \frac{\|v\|}{\|w\|}$$

is well-defined and (uniformly) continuous on the domain  $(v, w) \in \hat{P}$  and  $g \in \hat{V}$ . In particular, there exists  $\hat{\alpha} = \hat{\alpha}(\nu_\infty, n) > 0$  such that

$$(10.22) \quad -\log \frac{\|\Pi_{gv} gw\|}{\|gv\|} \frac{\|v\|}{\|w\|} \leq -\log \frac{\|\Pi_{fu} fz\|}{\|fu\|} \frac{\|u\|}{\|z\|} + \log 2$$

whenever  $d(v, u) < \hat{\alpha}$  and  $d(z, w) < \hat{\alpha}$  and  $d(g, f) < \hat{\alpha}$ . Reducing  $\hat{\alpha}$  if necessary, depending only on  $\nu_\infty$  and  $n$ , we may also suppose that

$$(10.23) \quad d(z, w) < \hat{\alpha} \Rightarrow -\log \frac{\|z^\perp\|}{\|z\|} \leq -\log \frac{\|w^\perp\|}{\|w\|} + \log 2.$$

Fix points  $(v_1, w_1), \dots, (v_l, w_l) \in \hat{P}$  such that the balls of radius  $\hat{\rho}$  around these points cover  $\hat{P}$ . For each  $(v, w) \in \hat{P}$  choose  $j \in \{1, \dots, l\}$  such that  $v \in B(v_j, \hat{\alpha})$  and  $w \in B(w_j, \hat{\alpha})$  and then define  $\hat{\mathcal{E}}_{k,0} = \hat{\mathcal{E}}_{k,0}(\nu_\infty, \delta, n, v, w) \subset G$  by

$$(10.24) \quad \hat{\mathcal{E}}_{k,0} = [\hat{\alpha}\text{-neighborhood of } \mathcal{E}_0(\nu_\infty, \delta, n, w_j^\perp)] \cap \text{supp } \nu_k^{(n)},$$

where  $\mathcal{E}_0(\nu_\infty, \delta, n, w_j^\perp)$  is given by Proposition 5.5. Since  $\nu_k^{(n)} \rightarrow \nu_\infty^{(n)}$  in the weak\* topology, the limit inferior of the  $\nu_k^{(n)}$ -measure of (10.24) as  $k \rightarrow \infty$  is greater than or equal to

$$\nu_\infty^{(n)}(\mathcal{E}_0(\nu_\infty, \delta, n, w_j^\perp)) > 1 - \delta$$

for every  $j = 1, \dots, l$ . In particular, there is  $k_1'' = k_1''(\nu_\infty, \delta, n) \in \mathbb{N}$  such that

$$(10.25) \quad \nu_k^{(n)}(\hat{\mathcal{E}}_{k,0}) > 1 - \delta \text{ for every } k \geq k_1'' \text{ and } (v, w) \in \hat{P}.$$

Given  $x \neq x'$  take  $w = v - v'$  to be the difference between unit vectors  $v$  and  $v'$  in the directions of  $x$  and  $x'$ , respectively. Just as in (8.22),

$$(10.26) \quad -\log d(gx, gx') \leq -\log d(x, x') - \log \frac{\|\Pi_{gv}gw\|}{\|gv\|} \frac{\|v\|}{\|w\|}$$

for any  $g \in G$  such that  $\|gv\| \geq \|gv'\|$  (the case  $\|gv\| \leq \|gv'\|$  is analogous, reversing the roles of  $x$  and  $x'$ ). The assumption  $x \in E_1(\varepsilon_1)$  implies that

$$(10.27) \quad \frac{\|v^\perp\|}{\|v\|} \leq d(x, E) \leq \varepsilon_1 < \hat{\rho}_1.$$

Now assume that  $-\log \text{VA}_1(x, x') < \tilde{\theta}_1 + \tilde{\kappa}_1 n$ . Then, using (8.19) and (8.23),

$$(10.28) \quad \frac{\|w^\perp\|}{\|w\|} > \frac{1}{4} \text{VA}_1(x, x') \geq \frac{1}{4} e^{-\tilde{\theta}_1 - \tilde{\kappa}_1 n} > 2\hat{\rho}_1.$$

Thus,  $(v, w) \in \hat{P}$ . Then define

$$(10.29) \quad \mathcal{E}_{k,1}'' = \hat{\mathcal{E}}_{k,0}(\nu_x, \delta, n, v, w).$$

It follows from (10.25) that  $\nu_k^{(n)}((\mathcal{E}_{k,1}'')^c) < \delta$  for every  $k \geq k_1''$ .

Take  $u = v_j$  and  $z = w_j$  as in (10.24). By definition,  $(u, z) \in \hat{P}$  and  $d(u, v) < \hat{\alpha}$  and  $d(z, w) < \hat{\alpha}$ . Let  $g \in \mathcal{E}_{k,1}''$  and  $k \geq \hat{k}_1$ . The definitions (10.24) and (10.29) imply that there exists  $f \in \mathcal{E}_0(\nu_x, \delta, n, z^\perp) \subset \text{supp } \nu_x^{(n)}$  such that  $d(g, f) < \hat{\alpha}$ . Replacing  $z, w$  and  $g$  with  $u, z$  and  $f$ , respectively, in (8.24) and (8.25), we find that

$$\frac{\|fu\|}{\|u\|} \geq \frac{1}{2} \frac{\|fu^E\|}{\|u^E\|} \quad \text{and} \quad \|\Pi_{fu}fz\| \geq \frac{1}{2} \|f^\perp z^\perp\|.$$

Combining this with (10.26) and (10.22), we obtain

$$(10.30) \quad \begin{aligned} -\log d(gx, gx') &\leq -\log d(x, x') + \log 2 - \log \frac{\|\Pi_{fu}fz\|}{\|fu\|} \frac{\|u\|}{\|z\|} \\ &\leq -\log d(x, x') + \log 8 - \log \frac{\|f^\perp z^\perp\|}{\|z\|} \frac{\|u^E\|}{\|fu^E\|}. \end{aligned}$$

Conditions (10.23) and (10.28) give that

$$-\log \frac{\|z^\perp\|}{\|z\|} \leq -\log \frac{\|w^\perp\|}{\|w\|} + \log 2 \leq -\log \text{VA}_1(x, x') + \log 8.$$

Substituting this in (10.30), we find that

$$\begin{aligned} -\log d(gx, gx') &\leq -\log d(x, x') - \log \text{VA}_1(x, x') + \log 64 - \log \frac{\|f^\perp z^\perp\|}{\|z^\perp\|} \frac{\|u^E\|}{\|fu^E\|} \\ &= -\log d(x, x') - \log \text{VA}_1(x, x') + \log 64 - \log \frac{\|Df_{u^E}^\perp z^\perp\|}{\|z^\perp\|}. \end{aligned}$$

By part (1) of Proposition 5.5 and the choice of  $\hat{N}_1 \in \mathbb{N}$  in (8.19), this implies that

$$\begin{aligned} -\log d(gx, gx') &\leq -\log d(x, x') - \log \text{VA}_1(x, x') + 5 - 2\hat{\kappa}_1 n \\ &\leq -\log d(x, x') - \log \text{VA}_1(x, x') - \hat{\kappa}_1 n, \end{aligned}$$

as claimed.  $\square$

Now we deduce the following extension of Proposition 8.3:

**Lemma 10.5.** *For  $\delta > 0$  and  $n \geq N_1$  and  $\omega > 0$  there is  $k_1''' = k_1'''(\nu_\infty, \delta, n, \omega) \in \mathbb{N}$  and for  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $\text{VA}_1(x, x') \geq \omega$  there is  $\mathcal{E}_{k,1}''' = \mathcal{E}_{k,1}'''(\nu_\infty, \delta, n, x, x', \omega) \subset \text{supp } \nu_k^{(n)}$  such that  $\nu_k^{(n)}((\mathcal{E}_{k,1}''')^c) < \delta$  and*

$$(10.31) \quad -\log \psi_1(gx, gx'; \omega) \leq -\log \psi_1(x, x'; \omega) - \kappa_1' n.$$

for any  $g \in \mathcal{E}_{k,1}'''$  and  $k \geq k_1'''$ .

*Proof.* Fix  $\delta > 0$  and  $n \geq N_1$  and  $\omega > 0$ . Recall that

$$N_1 > \max \{ \tilde{N}_1(\nu_\infty, \delta/2), \hat{N}_1(\nu_\infty, \delta/2), 2\tilde{\theta}_1/(\gamma\hat{\kappa}_1) \},$$

by (8.26). Define

$$(10.32) \quad \begin{aligned} k_1''' &= \max \{ k_1'(\nu_\infty, \delta/2, n, \omega), k_1''(\nu_\infty, \delta/2, n) \} \\ &\text{and } \mathcal{E}_{k,1}''' = \mathcal{E}_{k,1}'(\nu_\infty, \delta/2, n, x, x', \omega) \cap \mathcal{E}_{k,1}''(\nu_\infty, \delta/2, x, x', n). \end{aligned}$$

By construction,  $\mathcal{E}_{k,1}'''$  is contained in the support of  $\nu_k^{(n)}$  and  $\nu_k^{(n)}((\mathcal{E}_{k,1}''')^c) < \delta$ .

Consider  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $\text{VA}_1(x, x') \geq \omega$ . As observed in (9.9), the latter implies that  $\text{SVA}_1(x, x'; \omega) = \text{VA}_1(x, x')$  and  $\text{SVA}_1(gx, gx'; \omega) = \text{VA}_1(gx, gx')$ . Thus, the claim (10.31) may be rewritten as

$$(10.33) \quad -\log \text{VP}_1(gx, gx'; \omega) \leq -\log \text{VP}_1(x, x'; \omega) - \kappa_1' n.$$

Let  $g \in \mathcal{E}_{k,1}'''$  and  $k \geq k_1'''$ . Suppose first that  $-\log \text{VA}_1(x, x') \geq \tilde{\theta}_1 + \tilde{\kappa}_1 n$ . Then, by Lemma 10.3,

$$(10.34) \quad -\log \text{VA}_1(gx, gx') \leq -\log \text{VA}_1(x, x') - \tilde{\kappa}_1 n.$$

Substituting (10.34) and (8.2) in the definition (8.13), we find that

$$(10.35) \quad \begin{aligned} -\log \text{VP}_1(gx, gx') &\leq -\log \text{VA}_1(x, x') - \tilde{\kappa}_1 n - \gamma_1 \log d(x, x') + \gamma_1 A n \\ &\leq -\log \text{VP}_1(x, x') - \frac{\tilde{\kappa}_1}{2} n \end{aligned}$$

Now suppose that  $-\log \text{VA}_1(x, x') \leq \tilde{\theta}_1 + \tilde{\kappa}_1 n$ . In this case, Lemma 10.3 yields

$$(10.36) \quad -\log \text{VA}_1(gx, gx') \leq \tilde{\theta}_1,$$

whereas Lemma 10.4 asserts that

$$(10.37) \quad -\log d(gx, gx') \leq -\log d(x, x') - \log \text{VA}_1(x, x') - \hat{\kappa}_1 n.$$

Substituting (10.36) and (10.37) in the definition (8.13), we obtain

$$\begin{aligned} -\log \text{VP}_1(gx, gx') &\leq \tilde{\theta}_1 - \gamma_1 \log d(x, x') - \gamma_1 \log \text{VA}_1(x, x') - \gamma_1 \hat{\kappa}_1 n \\ &\leq -\log \text{VP}_1(x, x') + \tilde{\theta}_1 + (1 - \gamma_1) \log \text{VA}_1(x, x') - \gamma_1 \hat{\kappa}_1 n. \end{aligned}$$

Since  $\text{VA}_1(x, x') \leq 1$ ,  $\gamma_1 \leq 1$ , and  $n \geq N_1 \geq 2\tilde{\theta}_1/(\gamma_1\hat{\kappa}_1)$ , it follows that

$$(10.38) \quad \begin{aligned} -\log \text{VP}_1(gx, gx') &\leq -\log \text{VP}_1(x, x') + \tilde{\theta}_1 - \gamma_1 \hat{\kappa}_1 n \\ &\leq -\log \text{VP}_1(x, x') - \frac{\gamma_1 \hat{\kappa}_1}{2} n \end{aligned}$$

In view of choice of the constant  $\kappa_1'$  in (8.26), the claim (10.33) is contained in (10.35) and (10.38).  $\square$

Let us go back to proving Proposition 10.1. Define

$$(10.39) \quad k_1 = \max\{\tilde{k}_1, \hat{k}_1, k'_1, k''_1, k'''_1\}.$$

By construction,  $k_1$  depends only on  $\nu_\infty$ ,  $\delta$ ,  $n$  and  $\omega$ . Part (i) of the proposition is an immediate consequence of Lemma 10.2. To prove part (ii), consider the set  $\mathcal{E}'''_{k,1} = \mathcal{E}'''_{k,1}(\nu_\infty, \delta, n, x, x', \omega)$  given by Lemma 10.5. Then  $\nu_k^{(n)}((\mathcal{E}'''_{k,1})^c) < \delta$  and

$$(10.40) \quad -\log \psi_1(gx, gx'; \omega) \leq -\log \psi_1(x, x'; \omega) - \kappa'_1 n \leq -\log \psi_1(x, x'; \omega)$$

for every  $g \in \mathcal{E}'''_{k,1}$  and  $k \geq k'''_1$ . By (9.2), this implies that

$$(10.41) \quad \Psi_1(gx, gx'; \omega) \leq \Psi(x, x'; \omega) \text{ for every } g \in \mathcal{E}'''_{k,1}.$$

Integrating (10.41) over  $\mathcal{E}'''_{k,1}$  and (10.1) over the complement, we obtain the estimate in part (ii).

Next, recall that we took  $\kappa'''_1 = \kappa'_1/2$ ,  $c = e^{-2\kappa'''_1 n}$ , and  $\varepsilon'''_1 > 0$  such that (9.24) holds:

$$-\log \psi_1(x, x'; \omega) = -\log \psi_1(x, x') \geq \Omega_1/\sqrt{c}$$

for any  $x \neq x'$  in  $E_1(\varepsilon'''_1)$ . Then, by (9.3) and the first inequality in (10.40),

$$(10.42) \quad \Psi_1(gx, gx'; \omega) \leq \Psi_1(x, x'; \omega) + \log \sqrt{c} = \Psi_1(x, x'; \omega) - \kappa'''_1 n$$

for every  $g \in \mathcal{E}'''_{k,1}$ . Integrating (10.42) over  $\mathcal{E}'''_{k,1}$  and (10.1) over the complement, we obtain the estimate in part (iii) of the proposition. This completes the proof of Proposition 10.1.  $\square$

It is clear from the statements of Lemma 10.2, Lemma 10.3, and Lemma 10.5 that  $\tilde{k}_1$ ,  $k'_1$ , and  $k'''_1$  may be taken to increase to  $\infty$  when  $\omega$  decreases to zero and  $\nu_\infty$ ,  $\delta$ ,  $n$  remain fixed. Then the same is true about the map  $\omega \mapsto k_1(\nu_\infty, \delta, n, \omega)$  in (10.39). Hence, we may find  $\omega_{k,1} = \omega_{k,1}(\nu_\infty, \delta, n) > 0$  such that

$$(10.43) \quad (\omega_{k,1})_k \text{ decreases to } 0 \text{ and } k \geq k_1(\nu_\infty, \delta, n, \omega_{k,1})$$

for every large  $k \in \mathbb{N}$ . For instance,  $\omega_{k,1} = 2 \inf\{\omega > 0 : k_1(\nu_\infty, \delta, n, \omega) \leq k\}$ . Fix  $\tilde{k}_1 = \tilde{k}_1(\nu_\infty, \delta, n) \in \mathbb{N}$  such that (10.43) holds for all  $k \geq \tilde{k}_1$ . Define

$$(10.44) \quad \psi_{k,1}(x, x') = \psi_1(x, x'; \omega_{k,1}) \text{ and } \Psi_{k,1}(x, x') = \Psi_1(x, x'; \omega_{k,1}).$$

Then the following statement is contained in Proposition 10.1:

**Corollary 10.6.** *For any  $\delta > 0$ ,  $n \geq N_1$ , and  $k \geq \tilde{k}_1$ :*

(i) *For any  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $\Psi_{k,1}(x, x') > \log \Omega_1$ ,*

$$\int_G \Psi_{k,1}(gx, gx') d\nu_k^{(n)}(g) \leq \Psi_{k,1}(x, x') + C'_1 n.$$

(ii) *For any  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $\Psi_{k,1}(x, x') > \log \Omega_1$  and  $\text{VA}_1(x, x') \geq \omega_{k,1}$ ,*

$$\int_G \Psi_{k,1}(gx, gx') d\nu_k^{(n)}(g) \leq \Psi_{k,1}(x, x') + C'_1 \delta n.$$

(iii) *For any  $x \neq x'$  in  $E_1(\varepsilon'''_1)$  with  $\text{VA}_1(x, x') \geq \omega_{k,1}$ ,*

$$\int_G \Psi_{k,1}(gx, gx') d\nu_k^{(n)}(g) \leq \Psi_{k,1}(x, x') - (\kappa'_1 - C'_1 \delta) n.$$

## 11. RECOUPLING AND CONCLUSION

Here we resolve the difficulty arising out of the discontinuity of the Margulis function  $\Psi_{k,1}$ . The main issue is that the inequality in Lemma 10.2 may not hold when trajectories leave the cut-off region, that is, when  $\Psi_{k,1}(x, x'; n) = \log \Omega_1$  but  $\Psi(gx, gx'; n) \neq \log \Omega_1$  for some  $g$  in the support of  $\nu_k^{(n)}$ . That may cause the analogue of (6.26) to fail at such points, which is catastrophic for our proof.

From the form of the cut-off, that problem can only happen if both  $x$  and  $x'$  are outside  $E_1(2\varepsilon_1'')$  and at least one of  $gx$  or  $gx'$  is inside  $E_1(2\varepsilon_1'')$ . This is a relatively rare occurrence: for instance, Proposition 5.5 shows that, for every  $x$  not too far from  $E$ , the image  $gx$  is further away from  $E$  for the majority of  $g$  in the support of  $\nu_x^{(n)}$ . However, we still need to handle those rare cases where at least one of the points  $gx$  or  $gx'$  is in  $E_1(2\varepsilon_1'')$ .

The idea is to modify the dynamics on the space of pairs, more precisely the self-coupling  $\tilde{\mathcal{T}}_{k,1}$  of the Markov operator  $\mathcal{T}_{k,1}$ , to allow for the points  $x$  and  $x'$  to move in a more independent way: instead of the diagonal embedding, we will consider couplings supported on pairs  $(u, u')$  such that if one of the components is in  $E_1(2\varepsilon_1'')$  then the other lies outside  $E_1(\varepsilon_1')$ . Then  $\Psi_{k,1}(u, u')$  remains bounded above by a constant, which avoids the catastrophe.

**11.1. Recoupling.** As in Section 7.2, for each  $k \in \mathbb{N}$  let

$$\mathcal{T}_{k,1} : \mathcal{B}(E_1(\varepsilon_1)) \rightarrow \mathcal{B}(E_1(\varepsilon_1)), \quad \mathcal{T}_{k,1}\varphi(x) = \int_{E_1(\varepsilon_1)} \varphi(y) d\sigma_{k,1,x}(y)$$

be a continuous Markov operator adapted to  $(\nu_k^{(n)}, E_1(\varepsilon_1))$ , and let  $\eta_{k,1}$  be a  $\mathcal{T}_{k,1}$ -invariant probability measure converging, as  $k \rightarrow \infty$ , to a probability measure  $\eta_{\infty,1}$  such that  $\eta_{\infty,1}(E) > 0$ . Recall that  $\varepsilon_1 > 0$  was chosen small enough that

$$(11.1) \quad \eta_{\infty,1}(E_1(\varepsilon_1) \setminus E_1) < \frac{1}{10} \eta_{\infty,1}(E_1).$$

Then, for every  $k$  sufficiently large,

$$(11.2) \quad \eta_{k,1}(E_1(\varepsilon_1, \varepsilon_1''')) < \frac{2}{10} \eta_{k,1}(E_1(\varepsilon_1)).$$

Consider  $X = X' = E_1(\varepsilon_1)$  and  $Y = Y' = E_1(\varepsilon_1)$ , and  $\eta_y = \eta'_y = \sigma_{k,1,y}$  for every  $y \in Y$ . Moreover, let  $K = \text{Diag}_1$  be the diagonal of  $E_1(\varepsilon_1)^2$ . Clearly,  $K(x) = K'(x) = \{x\}$  for every  $x \in X$ . Since the  $\sigma_{k,1,y}$  are non-atomic measures, it follows that  $\sigma_{k,1,y}(K(x')) = \sigma_{k,1,y}(K'(x)) = \{0\}$  for every  $x, x'$  and  $y$ . Thus (6.10) holds in this case, and so we may use Proposition 6.9 to find a continuous family

$$\{\tilde{\theta}_{k,1,x,x'} : (x, x') \in E_1(\varepsilon_1)^2\}$$

of generic probability measures on  $E_1(\varepsilon_1)^2$  such that each  $\tilde{\theta}_{k,1,x,x'}$  is a coupling of  $\sigma_{k,1,x}$  and  $\sigma_{k,1,x'}$  vanishing on a uniform neighborhood of the diagonal.

Let  $\nu_{k,1,x}^{(n)}$  and  $\nu_{k,1,x,x'}^{(n)}$  denote the push-forwards of  $\nu_k^{(n)}$  under the maps  $G \rightarrow P$ ,  $g \mapsto gx$  and  $G \rightarrow P \times P$ ,  $g \mapsto (gx, gx')$ , respectively. Let  $\tilde{\omega} : E_1(\varepsilon_1)^2 \rightarrow [0, 1]$  be a continuous function such that  $\tilde{\omega}(x, x') = 0$  if  $x$  and  $x'$  are both in  $E_1(\tilde{\varepsilon}_1)$ , and  $\tilde{\omega}(x, x') = 1$  if either point is outside  $E_1(2\tilde{\varepsilon}_1)$ . Then

$$(11.3) \quad \tilde{\sigma}_{k,1,x,x'} = (1 - \tilde{\omega}(x, x')) \nu_{k,1,x,x'}^{(n)} + \tilde{\omega}(x, x') \tilde{\theta}_{k,1,x,x'}$$



is a coupling of  $\sigma_{k,1,x}$  and  $\sigma_{k,1,x'}$  depending continuously on  $(x, x')$ , and so

$$\tilde{\mathcal{T}}_{k,1} : \mathcal{B}(E_1(\varepsilon_1)^2) \rightarrow \mathcal{B}(E_1(\varepsilon_1)^2), \quad \tilde{\mathcal{T}}_{k,1}\varphi(x, x') = \int_{E_1(\varepsilon_1)^2} \varphi(y, y') d\tilde{\sigma}_{k,1,x,x'}(y, y'),$$

is a continuous self-coupling of  $\mathcal{T}_{k,1}$ . Since  $\nu_k^{(n)}$  and  $\tilde{\theta}_{k,1,x,x'}$  are generic measures, so is  $\tilde{\sigma}_{k,1,x,x'}$ . We are going to modify these Markov operators on the *recoupling region*  $E_1(2\tilde{\varepsilon}_1, \varepsilon_1'')^2$  as follows. See Figure 8.

For  $x \in E_1(2\tilde{\varepsilon}_1, \varepsilon_1'')$ , it follows from (9.16) that the subset of  $g \in \text{supp } \nu_k^{(n)}$  such that  $gx \in E_1(2\varepsilon_1')$  is disjoint from the set  $\mathcal{D}_k(x)$  given by Corollary 5.7. Hence,

$$(11.4) \quad \nu_{k,1,x}^{(n)}(E_1(2\varepsilon_1')) \leq \nu_k^{(n)}(\mathcal{D}_k(x)^c) < \delta \text{ for every } x \in E_1(2\tilde{\varepsilon}_1, \varepsilon_1'').$$

Keep in mind that  $\sigma_{k,1,x} = \nu_{k,1,x}^{(n)}$  if  $x$  is in the  $\nu_k^{(n)}$ -core of  $E_1(\varepsilon_1)$ .

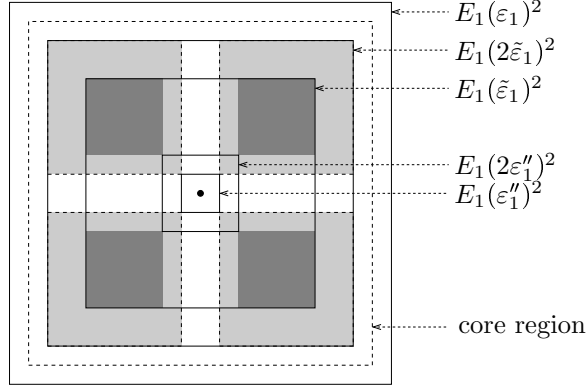


FIGURE 8. Illustrating the recoupling of the Markov operators. The black dot at the center marks the point  $(E, E)$ . The dashed lined corresponds to the boundary between the  $\nu_k^{(n)}$ -core and the  $\nu_k^{(n)}$ -border of  $E_1(\varepsilon_1)$ . The shaded area is the recoupling region. On the complement, marked in white,  $\hat{\sigma}_{k,1,x,x'} = \tilde{\sigma}_{k,1,x,x'}$ . On the dark gray area  $\tilde{\sigma}_{k,1,x,x'}$  is replaced with  $\zeta_{k,1,x,x'}$ , and on the light gray area we interpolate between the two.

Consider  $X = X' = E_1(\varepsilon_r)$ ,  $Y = Y' = E_1(2\tilde{\varepsilon}_1, \varepsilon_1'')^2$ , and  $\eta_y = \eta'_y = \sigma_{k,1,y}$  for every  $y \in Y$ . Moreover, let  $K = E_1(\varepsilon_1')^2 \cup \text{Diag}_1$ . Then

$$K(x') = E_1(\varepsilon_1') \cup \{x'\} \text{ and } K'(x) = E_1(\varepsilon_1') \cup \{x\}$$

for every  $(x, x') \in E_1(\varepsilon_1)^2$ . Since the  $\sigma_{k,1,y}$  are non-atomic measures, it follows from (11.4) that  $\sigma_{k,1,y}(K(x'))$  and  $\sigma_{k,1,y}(K'(x))$  are less than  $\delta < 1/2$  for every  $x, x'$  and  $y$ . This ensures that (6.10) holds in this case, and so we may use Proposition 6.9 to find a continuous family

$$\{\zeta_{k,1,x,x'} : (x, x') \in E_1(2\tilde{\varepsilon}_1, \varepsilon_1'')^2\}$$

of generic probability measures on  $E_1(\varepsilon_1)^2$  such that every  $\zeta_{k,1,x,x'}$  is a coupling of  $\sigma_{k,1,x}$  and  $\sigma_{k,1,x'}$  which vanishes on a uniform neighborhood of the diagonal and satisfies

$$(11.5) \quad \zeta_{k,1,x,x'}(E_1(\varepsilon_1')^2) = 0.$$

Fix a continuous function  $\tau : E_1(\varepsilon_1)^2 \rightarrow [0, 1]$  such that  $\tau \equiv 1$  on  $E_1(\tilde{\varepsilon}_1, 2\varepsilon_1'')^2$  and  $\tau \equiv 0$  on the complement of  $E_1(2\tilde{\varepsilon}_1, \varepsilon_1'')^2$ . Check Figure 8. Then define

$$(11.6) \quad \hat{\sigma}_{k,1,x,x'} = (1 - \tau(x, x')) \tilde{\sigma}_{k,1,x,x'} + \tau(x, x') \zeta_{k,1,x,x'}$$

for every  $(x, x') \in E_1(\varepsilon_1)^2$ . It is clear that  $\hat{\sigma}_{k,1,x,x'}$  is a coupling of  $\sigma_{k,1,x}$  and  $\sigma_{k,1,x'}$  depending continuously on  $(x, x')$ . Thus

$$\hat{\mathcal{T}}_{k,1} : \mathcal{B}(E_1(\varepsilon_1)^2) \rightarrow \mathcal{B}(E_1(\varepsilon_1)^2), \quad \hat{\mathcal{T}}_{k,1}\varphi(x, x') = \int_{E_1(\varepsilon_1)^2} \varphi(y, y') d\hat{\sigma}_{k,1,x,x'}(y, y')$$

is also a continuous self-coupling of  $\mathcal{T}_{k,1}$ . Moreover,  $\hat{\sigma}_{k,1,x,x'}$  is a generic measure, since  $\tilde{\sigma}_{k,1,x,x'}$  and  $\zeta_{k,1,x,x'}$  are generic, and it coincides with  $\tilde{\sigma}_{k,1,x,x'}$  outside the *recoupling region*  $E_1(2\tilde{\varepsilon}_1, \varepsilon_1'')^2$ .

**Lemma 11.1.** *Let  $(x, x') \in E_1(\varepsilon_1)^2$  be such that*

- (a) *either at least one of the points  $x$  or  $x'$  is in the  $\nu_k^{(n)}$ -border of  $E_1(\varepsilon_1)$ ,*
- (b) *or both  $x$  and  $x'$  are in the  $\nu_k^{(n)}$ -core of  $E_1(\varepsilon_1)$  but outside  $E_1(2\varepsilon_1')$ .*

*Then  $\hat{\sigma}_{k,1,x,x'}(E_1(\varepsilon_1')^2) = 0$  and so  $\hat{\mathcal{T}}_{k,1}\Psi_{k,1}(x, x') \leq \log \Omega_1 + \log 2$ .*

*Proof.* Let us begin by proving the claim that  $\hat{\sigma}_{k,1,x,x'}$  vanishes on  $E_1(\varepsilon_1')^2$ . If  $x$  is in the  $\nu_k^{(n)}$ -border of  $E_1(\varepsilon_1)$  then, using (9.17),

$$\hat{\sigma}_{k,1,x,x'}(E_1(\varepsilon_1')^2) \leq \sigma_{k,1,x}(E_1(\varepsilon_1')) \leq \sigma_{k,1,x} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_1) \right) = 0.$$

The same argument applies when  $x'$  is in the  $\nu_k^{(n)}$ -border of  $E_1(\varepsilon_1)$ . This settles the claim in case (a). Now let  $x$  and  $x'$  be as in (b). Keep in mind that  $\sigma_{k,1,x} = \nu_{k,1,x}^{(n)}$  and  $\sigma_{k,1,x'} = \nu_{k,1,x'}^{(n)}$ . By (11.5),  $\zeta_{k,1,x,x'}$  vanishes on  $E_1(\varepsilon_1')^2$ , and so (11.6) gives that

$$\hat{\sigma}_{k,1,x,x'}(E_1(\varepsilon_1')^2) = (1 - \tau(x, x')) \tilde{\sigma}_{k,1,x,x'}(E_1(\varepsilon_1')^2).$$

If  $x$  and  $x'$  are both in  $E_1(\tilde{\varepsilon}_1)$  then  $\tau(x, x') = 1$ , and the claim follows. When  $x \notin E_1(\tilde{\varepsilon}_1)$  we get from (9.18) that  $\nu_{k,1,x}^{(n)}(E_1(\varepsilon_1')) = 0$ . Then

$$\tilde{\sigma}_{k,1,x,x'}(E_1(\varepsilon_1')^2) \leq \sigma_{k,1,x}(E_1(\varepsilon_1')) = \nu_{k,1,x}^{(n)}(E_1(\varepsilon_1')) = 0.$$

The case when  $x' \notin E_1(\tilde{\varepsilon}_1)$  is analogous. We have shown that  $\hat{\sigma}_{k,1,x,x'}(E_1(\varepsilon_1')^2) = 0$  also in case (b).

By Lemma 9.5, it follows that  $\Psi_{k,1}(u, u') \leq \log \Omega_1 + \log 2$  for  $\hat{\sigma}_{k,1,x,x'}$ -almost every  $(u, u') \in E_1(\varepsilon_1)^2$ . Integrating with respect to  $\hat{\sigma}_{k,1,x,x'}$  we immediately get that  $\hat{\mathcal{T}}_{k,1}\Psi_{k,1}(x, x') \leq \log \Omega_1 + \log 2$ .  $\square$

**Proposition 11.2.** *There exist  $\kappa_1''' = \kappa_1'''(\nu_{\mathcal{C}}) > 0$  and  $C_1''' = C_1'''(\nu_{\mathcal{C}}) > 0$  such that given any  $\delta > 0$  and  $n \geq N_1$  the following holds for every  $k \geq \hat{k}_1$ :*

- (i) *For any  $x \neq x'$  in  $E_1(\varepsilon_1)$*

$$\hat{\mathcal{T}}_{k,1}\Psi_{k,1}(x, x') \leq \Psi_{k,1}(x, x') + C_1''' n.$$

- (ii) *For any  $x \neq x'$  in  $E_1(\varepsilon_1)$  with  $\text{VA}_1(x, x') \geq \omega_{k,1}$ ,*

$$\hat{\mathcal{T}}_{k,1}\Psi_{k,1}(x, x') \leq \Psi_{k,1}(x, x') + C_1'''(1 + \delta n).$$

- (iii) *For any  $x \neq x'$  in  $E_1(\varepsilon_1''')$  with  $\text{VA}_1(x, x') \geq \omega_{k,1}$ ,*

$$\hat{\mathcal{T}}_{k,1}\Psi_{k,1}(x, x') \leq \Psi_{k,1}(x, x') - (\kappa_1''' - C_1''' \delta) n.$$

*Proof.* Take  $\kappa_1''' = \kappa_1'$  and  $C_1''' = \max\{C_1', \log 2\}$ , and let  $k \geq \check{k}_1$ . We split the argument into four cases (check Figure 8).

First, suppose that both  $x$  and  $x'$  are in the  $\nu_k^{(n)}$ -core of  $E_1(\varepsilon_1)$ , and at least one of them is in  $E_1(\varepsilon_1'')$ . This is necessarily the case in the setting of (iii). In particular  $(x, x')$  is outside the cut-off region,

$$\Psi_{k,1}(x, x') = \log(\Omega_1 + \psi_{k,1}(x, x')^{-1}) > \log \Omega_1,$$

and there is no recoupling either:

$$\hat{\sigma}_{k,1,x,x'} = \nu_{k,1,x,x'}^{(n)}, \quad \hat{\mathcal{T}}_{k,1}\Psi_{k,1}(x, x') = \int_G \Psi_{k,1}(gx, gx') d\nu_k^{(n)}(g).$$

Hence the claims in (i), (ii) and (iii) are contained in Corollary 10.6.

Now suppose that both  $x$  and  $x'$  are in the  $\nu_k^{(n)}$ -core of  $E_1(\varepsilon_1)$  but outside  $E_1(\varepsilon_1'')$ , and at least one of them is in  $E_1(2\varepsilon_1'')$ . It is still true that  $(x, x')$  is outside the cut-off region, and so  $\Psi_{k,1}(x, x') > \log \Omega_1$ . Thus the estimates in Corollary 10.6 remain valid for

$$(11.7) \quad \int \Psi_{k,1} d\tilde{\sigma}_{k,1,x,x} = \int_G \Psi_{k,1}(gx, gx') d\nu_k^{(n)}(g).$$

However, this time  $(x, x')$  may be in the recoupling region. That is dealt with as follows. By (11.5), the measure  $\zeta_{k,1,x,x'}$  vanishes on  $E_1(\varepsilon_1')^2$ . Thus, by Lemma 9.5,

$$(11.8) \quad \int_{E_1(\varepsilon_1)^2} \Psi_{k,1} d\zeta_{k,1,x,x} \leq \log \Omega_1 + \log 2 \leq \Psi_{k,1}(x, x') + \log 2.$$

The claims (i) and (ii) follow because, by (11.6),  $\hat{\mathcal{T}}_{k,1}\Psi_{k,1}(x, x')$  is a convex combination of the integrals in (11.7) and (11.8).

Next suppose that both  $x$  and  $x'$  are in the  $\nu_k^{(n)}$ -core of  $E_1(\varepsilon_1)$  but outside  $E_1(2\varepsilon_1'')$ . This corresponds to case (b) of Lemma 11.1: claims (i) and (ii) are contained in the conclusion of that lemma. Finally, suppose that at least one of the points  $x$  and  $x'$  is in the  $\nu_k^{(n)}$ -border of  $E_1(\varepsilon_1)$ . This is precisely the situation in case (a) of Lemma 11.1, and so claims (i) and (ii) are again contained in the conclusion of that lemma.  $\square$

**11.2. Contradicting  $\dim E = 1$ .** We are going to use the following refinement of Lemma 6.21:

**Lemma 11.3.** *Let  $\mathcal{T} : \mathcal{B}(X) \rightarrow \mathcal{B}(X)$  be a Markov operator and  $\psi : X \rightarrow [0, \infty]$  be a measurable function. Suppose that there exist constants  $\kappa_A, \kappa_B', \kappa_B'' \in \mathbb{R}$  and pairwise disjoint sets  $A, B', B''$  such that  $X = A \cup B' \cup B''$  and*

- (i)  $\mathcal{T}\psi(x) \leq \psi(x) - \kappa_A$  for  $x \in A$ ,
- (ii)  $\mathcal{T}\psi(x) \leq \psi(x) + \kappa_B'$  for  $x \in B'$ ,
- (iii)  $\mathcal{T}\psi(x) \leq \psi(x) + \kappa_B''$  for  $x \in B''$ .

Let  $\hat{\eta}$  be a measure on  $X$  with  $\int \psi d\hat{\eta} < \infty$  and  $\int_X \mathcal{T}\psi(x) d\hat{\eta}(x) \geq \int_X \psi(x) d\hat{\eta}(x)$ . Then

$$(11.9) \quad \hat{\eta}(B'') \geq \frac{\kappa_A \hat{\eta}(X) - (\kappa_A + \kappa_B') \hat{\eta}(B')}{\kappa_A + \kappa_B''}.$$

*Proof.* Conditions (i) - (iii) imply

$$\begin{aligned} \int_X \psi(x) d\hat{\eta}(x) &\leq \int_X \mathcal{T}\psi(x) d\hat{\eta}(x) \\ &\leq \int_X \psi(x) d\hat{\eta}(x) - \kappa_A\theta(A) + \kappa'_B\theta(B') + \kappa''_B\theta(B''). \end{aligned}$$

Thus,  $-\kappa_A\hat{\eta}(A) + \kappa'_B\hat{\eta}(B') + \kappa''_B\hat{\eta}(B'') \geq 0$ , which implies (11.9).  $\square$

Take  $X = E_1(\varepsilon_1)^2$ ,  $\mathcal{T} = \hat{\mathcal{T}}_{k,1}$ ,  $\psi = \Psi_{k,1}$ ,  $\hat{\eta} = \hat{\eta}_{k,1}$ , and

$$(11.10) \quad A_k = \{(x, x') \in E_1(\varepsilon_1'')^2 : \text{VA}_1(x, x') > \omega_{k,1}\},$$

$$(11.11) \quad B' = \{(x, x') \in E_1(\varepsilon_1)^2 : d(x, E) > \varepsilon_1''' \text{ or } d(x', E) > \varepsilon_1'''\},$$

$$(11.12) \quad B''_k = \{(x, x') \in E_1(\varepsilon_1)^2 : \text{VA}_1(x, x') \leq \omega_{k,1}\}.$$

It is clear that  $A = A_k$  is disjoint from  $B = B' \cup B''_k$ , and their union is the whole  $E_1(\varepsilon_1)^2$ . The sets  $B'$  and  $B'' = B''_k$  are also disjoint if  $k$  is sufficiently large, because

$$\text{VA}_1(x, x') = d(x + x', E) \geq \max\{d(x, E), d(x', E)\}$$

is greater than  $\varepsilon_1'''$  whereas  $(\omega_{k,1})_k$  converges to zero when  $k \rightarrow \infty$ . Also,  $A \cup B' \cup B''$  is the whole  $E_1(\varepsilon_1)^2$ . Moreover, (7.2) implies that  $B'' = \emptyset$  when  $\dim E = 1$ .

Proposition 11.2 shows that, assuming that  $k$  is sufficiently large, the hypotheses of Lemma 11.3 are satisfied for these choices, with

$$\kappa_A = (\kappa_1''' - C_1''' \delta)n, \quad \kappa'_B = C_1'''(1 + \delta)n, \quad \text{and} \quad \kappa''_B = C_1'''n.$$

Take  $\delta > 0$  to be sufficiently small, depending on  $\nu_\infty$ , and  $n \in \mathbb{N}$  to be sufficiently large, depending on  $\nu_\infty$  and  $\delta$ , that

$$(11.13) \quad \kappa_A > 9\kappa''_B.$$

As in Proposition 6.23, we find a sequence  $(\hat{\eta}_{k,1,j})_j$  of probability measures on  $E_1(\varepsilon_1)^2$  converging to a  $\hat{\mathcal{T}}_{k,1}$ -invariant self-coupling  $\hat{\eta}_{k,1}$  of  $\eta_{k,1}$  such that every  $\hat{\eta}_{k,1,j}$  satisfies  $\int_{E_1(\varepsilon_1)^2} \Psi_{k,1} d\hat{\eta}_{k,1,j} < \infty$  and

$$\int_{E_1(\varepsilon_1)^2} \hat{\mathcal{T}}_{k,1} \Psi_{k,1}(x, x') d\hat{\eta}_{k,1,j}(x, x') \geq \int_{E_1(\varepsilon_1)^2} \Psi_{k,1}(x, x') d\hat{\eta}_{k,1,j}(x, x').$$

Applying Lemma 11.3 with  $\mathcal{T} = \hat{\mathcal{T}}_{k,1}$  and  $\hat{\eta} = \hat{\eta}_{k,1,j}$  we get that

$$(11.14) \quad \hat{\eta}_{k,1,j}(B'') \geq \frac{\kappa_A \hat{\eta}_{k,1,j}(E_1(\varepsilon_1)^2) - (\kappa_A + \kappa'_B) \hat{\eta}_{k,1,j}(B')}{\kappa_A + \kappa''_B}$$

for every  $j$ . Passing to the limit as  $j \rightarrow \infty$ , we conclude that

$$(11.15) \quad \hat{\eta}_{k,1}(B'') \geq \frac{\kappa_A \hat{\eta}_{k,1}(E_1(\varepsilon_1)^2) - (\kappa_A + \kappa'_B) \hat{\eta}_{k,1}(B')}{\kappa_A + \kappa''_B}.$$

Observe that  $\hat{\eta}_{k,1}(E_1(\varepsilon_1)^2) = \eta_{k,1}(E_1(\varepsilon_1))$  and, using (11.2),

$$(11.16) \quad \hat{\eta}_{k,1}(B') \leq 2\eta_{k,1}(E_1(\varepsilon_1, \varepsilon_1''')) < \frac{4}{10}\eta_{k,1}(E_1(\varepsilon_1)).$$

Thus, (11.13) and (11.15) imply

$$(11.17) \quad \hat{\eta}_{k,1}(B'') \geq \frac{\kappa_A - \frac{4}{10}(\kappa_A + \kappa'_B)}{\kappa_A + \kappa''_B} \eta_{k,1}(E_1(\varepsilon_1)) \geq \frac{5\kappa'_B}{\kappa_A + \kappa''_B} \eta_{k,1}(E_1(\varepsilon_1)) > 0.$$

When  $\dim E = 1$  this yields a contradiction, because  $B''$  is empty in that case. Thus  $\dim E \geq 2$ .

**11.3. Completing the first step.** Let us consider the map

$$\Sigma : E_1(\varepsilon_1)^2 \setminus \text{Diag}_1 \rightarrow \text{Gr}(2, d), \quad \Sigma(x, x') = x + x'.$$

We would like to define  $\eta_{k,2} = \Sigma_*(\tilde{\eta}_{k,1})$  but there is a problem in that  $\Sigma(x, x')$  is not defined on  $\text{Diag}_1$  and we cannot exclude the possibility that  $\tilde{\eta}_{k,1}$  is positive on the diagonal.

To by-pass this difficulty, we introduce the compact topological spaces

$$\begin{aligned} \mathcal{Y}_1 &= \{(x, x', y) \in \text{Gr}(1, d)^2 \times \text{Gr}(2, d) : x \subset y \text{ and } x' \subset y\} \\ \mathcal{Y}_1(\varepsilon) &= \{(x, x', y) \in \mathcal{Y}_1 : x, x' \in E_1(\varepsilon)\} \text{ for } \varepsilon > 0, \end{aligned}$$

together with the canonical projections

$$\begin{aligned} p_1 : \mathcal{Y}_1 &\rightarrow \text{Gr}(1, d)^2, & (x, x', y) &\mapsto (x, x') \\ p_2 : \mathcal{Y}_1 &\rightarrow \text{Gr}(2, d), & (x, x', y) &\mapsto y. \end{aligned}$$

For  $(x, x', y) \in \mathcal{Y}_1$ ,  $k \in \mathbb{N}$ , and  $n \in \mathbb{N}$ , denote by  $\nu_{k,1,x,x',y}^{(n)}$  the image of  $\nu_k^{(n)}$  under the diagonal action

$$G \rightarrow \mathcal{Y}_1, \quad (g \mapsto (gx, gx', gy)).$$

Clearly, each  $\nu_{k,1,x,x',y}^{(n)}$  is a lift of  $\nu_{k,1,x,x'}^{(n)}$  relative to  $p_1 : \mathcal{Y}_1 \rightarrow \text{Gr}(1, d)^2$ . The complement of the diagonal in  $\text{Gr}(1, d)^2$  embeds in  $\mathcal{Y}_1$  through

$$(x, x') \mapsto (x, x', x + x').$$

In particular, every measure  $\xi$  on  $\text{Gr}(1, d)^2$  that vanishes on the diagonal has a (unique) lift  $\tilde{\xi}$  to  $\mathcal{Y}_1$ .

From the relations (11.3) and (11.6), we see that

$$(11.18) \quad \hat{\sigma}_{k,1,x,x'} = (1 - \tilde{\omega}(x, x')) \nu_{k,1,x,x'}^{(n)} + \tilde{\omega}(x, x') \hat{\theta}_{k,1,x,x'}$$

where  $\tilde{\omega} : E_1(\varepsilon_1)^2 \rightarrow [0, 1]$  is a continuous function that vanishes identically on  $E_1(\varepsilon_1'')^2$ , and each  $\hat{\theta}_{k,1,x,x'}$  is a coupling of  $\sigma_{k,1,x}$  and  $\sigma_{k,1,x'}$  vanishing on a uniform neighborhood of the diagonal. In view of the previous remarks, it follows that the  $\hat{\sigma}_{k,1,x,x'}$  lift to probability measures

$$(11.19) \quad \check{\sigma}_{k,1,x,x',y} = (1 - \tilde{\omega}(x, x')) \nu_{k,1,x,x',y}^{(n)} + \tilde{\omega}(x, x') \check{\theta}_{k,1,x,x',y}$$

on  $\mathcal{Y}_1(\varepsilon_1)$ , where  $\check{\theta}_{k,1,x,x',y}$  is the unique lift of  $\hat{\theta}_{k,1,x,x'}$ . Since  $\nu^{(n)}$  and  $\hat{\theta}_{k,1,x,x'}$  are generic measures, so is  $\check{\sigma}_{k,1,x,x',y}$  for every  $(x, x', y) \in \mathcal{Y}_1(\varepsilon_1)$ .

It is clear that  $\nu_{k,x,x',y}^{(n)}$  varies continuously on  $\mathcal{Y}_1(\varepsilon_1)$  and, by uniqueness, so does  $\check{\theta}_{k,1,x,x',y}$ . Thus,

$$\check{\mathcal{T}}_{k,1} : \mathcal{B}(\mathcal{Y}_1(\varepsilon_1)) \rightarrow \mathcal{B}(\mathcal{Y}_1(\varepsilon_1)), \quad \check{\mathcal{T}}_{k,1}\Psi(x, x', y) = \int_{\mathcal{Y}_1(\varepsilon_1)} \Psi d\check{\sigma}_{k,1,x,x',y}.$$

defines a continuous Markov operator. From the definition we see that  $\check{\mathcal{T}}_{k,1}$  is a lift of  $\hat{\mathcal{T}}_{k,1}$ , in the sense that

$$\check{\mathcal{T}}_{k,1}(\psi \circ p_1) = \left( \hat{\mathcal{T}}_{k,1}\psi \right) \circ p_1 \text{ for every } \psi \in \mathcal{B}(\mathcal{Y}_1(\varepsilon_1)).$$

Thus, the construction in Proposition 6.23 can be applied simultaneously to the two operators, to yield a sequence of probability measures  $\check{\eta}_{k,1,j}$  converging to

a  $\tilde{\mathcal{T}}_{k,1}$ -invariant measure  $\check{\eta}_{k,1}$  on  $\mathcal{Y}_1(\varepsilon_1)$ , and whose projections down to  $E_1(\varepsilon_1)^2$  are self-couplings  $\hat{\eta}_{k,1,j}$  of the  $\eta_{k,1}$  vanishing on neighborhoods of the diagonal of  $E_1(\varepsilon_1)^2$ , and converging to the  $\tilde{\mathcal{T}}_{k,1}$ -invariant measure  $\hat{\eta}_{k,1}$ .

Next, define  $\eta_{k,2} = p_{2*}\check{\eta}_{k,1}$  and let  $\{d\check{\eta}_{k,1,v} : v \in p_2\mathcal{Y}_1(\varepsilon_1)\}$  be a disintegration of  $\check{\eta}_{k,1}$  with respect to the partition  $\{p_2^{-1}(v) : v \in p_2\mathcal{Y}_1(\varepsilon_1)\}$ . Then define

$$\begin{aligned} \mathcal{T}_{k,2} &: \mathcal{B}(p_2\mathcal{Y}_1(\varepsilon_1)) \rightarrow \mathcal{B}(p_2\mathcal{Y}_1(\varepsilon_1)), \\ \mathcal{T}_{k,2}\Phi(y) &= \int_{p_2^{-1}(y)} \tilde{\mathcal{T}}_{k,1}(\Phi \circ p_2)(x, x', y) d\check{\eta}_{k,1,y}(x, x'). \end{aligned}$$

Equivalently,  $\mathcal{T}_{k,2}\Phi(y) = \int_{p_2\mathcal{Y}_1(\varepsilon_1)} \Phi d\sigma_{k,2,y}$  with

$$(11.20) \quad \sigma_{k,2,y} = \int_{p_2^{-1}(y)} p_{2*}\check{\sigma}_{k,1,x,x',y} d\check{\eta}_{k,1,y}(x, x').$$

Let  $\mathcal{B}_k = p_2p_1^{-1}(B_k'') = \{x + x' : (x, x') \in B_k''\}$ , where  $B_k''$  is as in the previous section. Define also  $\eta_{k,2,j} = p_{2*}\check{\eta}_{k,1,j}$  for  $j \in \mathbb{N}$ . Then

$$\eta_{k,2,j}(\mathcal{B}_k) \geq \check{\eta}_{k,1,j}(p_1^{-1}(B_k'')) = \hat{\eta}_{k,1,j}(B_k'').$$

Passing to the limit as  $j \rightarrow \infty$  and arguing as in (11.14)–(11.17) we find that

$$\eta_{k,2}(\mathcal{B}_k) \geq \frac{5\kappa'_B}{\kappa_A + \kappa''_B} \eta_{k,1}(E_1(\varepsilon_1)).$$

Now, the definition (11.12) implies that  $\mathcal{B}_k$  converges to  $E_2$  as  $k \rightarrow \infty$ , because  $\omega_{k,1} \rightarrow 0$ . Thus, any accumulation point  $\eta_{\infty,2}$  of  $\eta_{k,2}$  must satisfy

$$(11.21) \quad \eta_{\infty,2}(E_2) \geq \frac{5\kappa'_B}{\kappa_A + \kappa''_B} \eta_{\infty,1}(E_1) > 0.$$

Take  $n_2 = n$  and  $\varepsilon_2 = \varepsilon_1''$ . Let  $\nu_{k,2,y}^{(n_2)}$  denote the push-forward of  $\nu_k^{(n_2)}$  under the map  $G \rightarrow \text{Gr}(2, d)$ ,  $g \mapsto gy$ .

**Lemma 11.4.**

- (i)  $\sigma_{k,2,y} = \nu_{k,2,y}^{(n_2)}$  for every  $y \in E_2(\varepsilon_2)$ .
- (ii)  $\sigma_{k,2,y} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_2(\varepsilon_2) \right) = 0$  for every  $y \notin E_2(\varepsilon_2)$
- (iii) The measure  $\eta_{k,2}$  is  $\mathcal{T}_{k,2}$ -invariant.

*Proof.* It is clear that  $\nu_{k,2,y}^{(n_2)}$  coincides with the push-forward of  $\nu_{k,1,x,x',y}^{(n_2)}$  under the projection  $p_2$ . Thus (11.19) gives that

$$p_{2*}\check{\sigma}_{k,1,x,x',y} = (1 - \check{\omega}(x, x')) \nu_{k,2,y}^{(n_2)} + \check{\omega}(x, x') p_{2*}\check{\theta}_{k,1,x,x',y},$$

and so,

$$\begin{aligned} \check{\sigma}_{k,2,y} &= \left( 1 - \int_{p_2^{-1}(y)} \check{\omega}(x, x') d\check{\eta}_{k,1,y}(x, x') \right) \nu_{k,2,y}^{(n_2)} \\ &\quad + \int_{p_2^{-1}(y)} \check{\omega}(x, x') p_{2*}\check{\theta}_{k,1,x,x',y} d\check{\eta}_{k,1,y}(x, x'), \end{aligned}$$

If  $y \in E_2(\varepsilon_2)$  then both  $x$  and  $x'$  are necessarily in  $E_1(\varepsilon_2)$ , by (7.1), in which case  $\check{\omega}(x, x') = 0$ . Then  $\sigma_{k,2,y} = \nu_{k,2,y}^{(n_2)}$ , as claimed in (i).

In view of the expression (11.20), to prove part (ii) it suffices to show that if  $y \notin E_2(\varepsilon_2)$  then

$$(11.22) \quad p_{2*} \check{\sigma}_{k,1,x,x',y} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_2(\varepsilon_2) \right) = 0$$

for any  $x, x' \subset y$ . If  $x$  and  $x'$  are both in  $E_1(\varepsilon_2)$  then

$$p_{2*} \check{\sigma}_{k,1,x,x',y} = p_{2*} \nu_{k,1,x,x',y}^{(n_2)} = \nu_{k,2,y}^{(n_2)}$$

and then the claim follows from Remark 6.26. From now on, we assume that one of the points,  $x$  say, is not in  $E_1(\varepsilon_2)$ . It follows from the definitions that

$$p_2^{-1} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_2(\varepsilon_2) \right) \subset \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_2)^2 \times \mathcal{X}_{\nu_k^{(n)}}^\# E_2(\varepsilon_2),$$

and so

$$(11.23) \quad \begin{aligned} p_{2*} \check{\sigma}_{k,1,x,x',y} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_2(\varepsilon_2) \right) &\leq \check{\sigma}_{k,1,x,x',y} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_2)^2 \times \mathcal{X}_{\nu_k^{(n)}}^\# E_2(\varepsilon_2) \right) \\ &\leq \hat{\sigma}_{k,1,x,x'} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_2)^2 \right) \\ &\leq \sigma_{k,1,x} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_2) \right). \end{aligned}$$

If  $x$  is in the  $\nu_k^{(n)}$ -border of  $E_1(\varepsilon_1)$  then

$$\sigma_{k,1,x} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_2) \right) \leq \sigma_{k,1,x} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_1) \right) = 0,$$

because the operator  $\mathcal{T}_{k,1}$  is adapted to  $(\nu_k^{(n)}, E_1(\varepsilon_1))$ . If  $x$  is in the  $\nu_k^{(n)}$ -core of  $E_1(\varepsilon_1)$  then Remark 6.26 gives that

$$\sigma_{k,1,x} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_2) \right) = \nu_{k,1,x}^{(n)} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_1(\varepsilon_2) \right) = 0.$$

Thus the right-hand side of (11.23) vanishes in either case. That completes the proof of (11.22) and of part (ii) of the lemma.

Finally, by definition,

$$\begin{aligned} &\int_{p_2 \mathcal{Y}_1(\varepsilon_1)} (\mathcal{T}_{k,2} \Phi) d\eta_{k,2} \\ &= \int_{p_2 \mathcal{Y}_1(\varepsilon_1)} \int_{p_2^{-1}(y)} \check{\mathcal{T}}_{k,1}(\Phi \circ p_2)(x, x', y) d\check{\eta}_{k,1,y}(x, x') d\eta_{k,2}(y) \\ &= \int_{\mathcal{Y}_1(\varepsilon_1)} \check{\mathcal{T}}_{k,1}(\Phi \circ p_2)(x, x', y) d\check{\eta}_{k,1}(x, x', y) \end{aligned}$$

for any  $\Phi \in \mathcal{B}(p_2 \mathcal{Y}_1(\varepsilon_1))$ . Since  $\check{\eta}_{k,1}$  is  $\check{\mathcal{T}}_{k,1}$ -invariant, this gives

$$\begin{aligned} \int_{p_2 \mathcal{Y}_1(\varepsilon_1)} (\mathcal{T}_{k,2} \Phi)(y) d\eta_{k,2}(y) &= \int_{\mathcal{Y}_1(\varepsilon_1)} (\Phi \circ p_2)(x, x', y) d\check{\eta}_{k,1}(x, x', y) \\ &= \int_{p_2 \mathcal{Y}_1(\varepsilon_1)} \Phi d\hat{\eta}_{k,2}, \end{aligned}$$

which proves claim (iii).  $\square$

Since the  $\check{\sigma}_{k,1,x,x',y}$  are generic measures and the projection  $p_2$  is algebraic, it follows readily from (11.20) and Remark 5.1 that every  $\sigma_{k,2,y}$  is a generic measure. Then, conclusions (i) and (ii) in Lemma 11.4 allow us to apply Propositions 6.15 and 6.25 to  $X = p_2\mathcal{Y}_1(\varepsilon_1)$  and  $U = E_2(\varepsilon_2)$ . In this way we get a continuous Markov operator adapted to  $(\nu_k^{(n_2)}, E_2(\varepsilon_2))$  which leaves invariant the restriction of  $\eta_{k,2} \mid E_2(\varepsilon_2)$ . Replace  $\mathcal{T}_{k,2}$  and  $\eta_{k,2}$  with this new Markov operator and invariant measure, respectively. This finishes the first step of the induction.

## Part IV. General step of the induction

### 12. PREPARING A MARGULIS FUNCTION

Let  $r > 1$  be fixed. Here we extend the construction in Section 8, to find a positive function  $\psi_r$  such that  $-\log \psi_r$  has some of the features of a Margulis function. The main result is Proposition 12.8, an extension of Proposition 8.1. Throughout,  $\delta > 0$  and  $n \in \mathbb{N}$  should be seen as free parameters, whose values are fixed at the end of the construction.

Recall that  $\mathcal{F}(r, d)$  denotes the set of flags  $F_1 \subset F_2 \subset \dots \subset F_{r-1} \subset F_r \subset \mathbb{R}^d$ , where each  $F_i$  has dimension  $i$ . Moreover,

$$E_r^\circ = \{(F_1, \dots, F_r) \in \mathcal{F}(r, d) : F_r \in E_r\} \text{ and,}$$

$$E_r^\circ(\varepsilon) = \{(F_1, \dots, F_r) \in \mathcal{F}(r, d) : F_r \in E_r(\varepsilon)\} \text{ for each } \varepsilon > 0.$$

We use  $x = (F_1, \dots, F_r)$  and  $x' = (F'_1, \dots, F'_r)$  to denote generic elements of  $\mathcal{F}(r, d)$ .

Most steps towards Proposition 12.8 are rather straightforward translations of the arguments in Section 8. One significant difference is that Lemma 8.6 no longer holds: for  $r > 1$  it is possible to have  $x$  and  $x'$  with  $F_r$  and  $F'_r$  arbitrarily close to the equator  $E$  without  $\text{VP}_r(x, x')$  getting close to zero. For that reason, we cannot take  $\psi_r = \text{VP}_r$ . This is dealt with in Section 12.3: instead, we define  $\psi_r$  inductively in terms of both  $\text{VP}_r$  and  $\psi_{r-1}$ .

**12.1. Vertical angle function.** Assume  $x, x' \in \mathcal{F}(r, d)$  to be such that  $F'_1 \not\subset F_r$ . By definition, the *great circle* through  $x$  and  $x'$  is the subset  $y = y(x, x')$  of  $\text{Gr}(r, d)$  defined by

$$y = \{\xi \in \text{Gr}(r, d) : F_{r-1} \subset \xi \subset F'_1 + F_r\}.$$

This is consistent with the case  $r = 1$ , as long as we follow the convention that  $F_0 = \{0\}$ . On the other hand, the great circle  $y$  depends on  $x$  through  $F_{r-1}$  and  $F_r$ , whereas it depends on  $x'$  through  $F'_1$  only. In particular,  $y(x, x')$  need not coincide with  $y(x', x)$  when  $r > 1$ . Related to this, the analogue of (8.14) is usually false when  $r > 1$ : the values of  $\text{VA}_r(x, x')$  and  $\text{VP}_r(x, x')$  that we define in the following may change when the roles of  $x$  and  $x'$  are exchanged.

The *vertical angle function*  $\text{VA}_r$  is defined by

$$(12.1) \quad \text{VA}_r(x, x') = \sup_{\xi \in y} d(\xi, E) = \sup_{\xi \in y} \sup_{u \in \xi} d(u, E).$$

Note that this is consistent with the case  $r = 1$ . Following the intuition from  $r = 1$ , we think of  $\text{VA}_r(x, x')$  as the angle between the great circle  $y$  and  $E_r$ . Indeed,

$$(12.2) \quad \text{VA}_r(x, x') = \sup_{u \in F'_1 + F_r} d(u, E) = d(F'_1 + F_r, E),$$

and so  $\text{VA}_r(x, x') = 0$  if and only if  $F'_1 + F_r \subset E$ , that is, if and only if  $y \subset E_r$ .

We are going to prove the following extension of Proposition 8.1 for  $r > 1$ :



**Proposition 12.1.** *There exist  $\kappa'_r = \kappa'_r(\nu_{\mathcal{O}}) > 0$  and  $C'_r = C'_r(\nu_{\mathcal{O}}) > 0$  and for any  $\delta > 0$  there exists  $N_r = N_r(\nu_{\mathcal{O}}, \delta)$  such that for every  $n \geq N_r$  there exists  $\rho'_r = \rho'_r(\nu_{\mathcal{O}}, \delta, n) > 0$  such that*

$$\int_G -\log \text{VP}_r(gx, gx') d\nu_{\mathcal{O}}^{(n)}(g) \leq -\log \text{VP}_r(x, x') - (\kappa'_r - C'_r \delta)n$$

for every  $x, x' \in E_r^\circ(\rho'_r)$  with  $F_1' \not\subset F_r$ .

Begin by noting that the properties (8.5) through (8.7) extend to  $r > 1$ . More precisely, it follows from (12.2) that

$$(12.3) \quad \text{VA}_r(x, x') \geq d(u, E) = \frac{\|u^\perp\|}{\|u\|}$$

for any  $u \in F_1' + F_r$  and  $x, x' \in \mathcal{F}(r, d)$  with  $F_1' \not\subset F_r$ . Moreover, if  $u \in F_1' + F_r$  realizes the supremum in the definition (12.2) then

$$(12.4) \quad \text{VA}_r(x, x') = d(u, E) = \frac{\|u^\perp\|}{\|u\|}.$$

Then, combining (12.3) with (12.4),

$$(12.5) \quad \text{VA}_r(gx, gx') \geq \frac{\|(gu)^\perp\|}{\|gu\|} = \text{VA}_r(x, x') \frac{\|(gu)^\perp\|}{\|gu\|} \frac{\|u\|}{\|u^\perp\|}$$

for any  $g \in G$ . By Remark 3.1, when  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$  this means that

$$(12.6) \quad \text{VA}_r(gx, gx') \geq \frac{\|g^\perp u^\perp\|}{\|gu\|} \geq \text{VA}_r(x, x') \frac{\|g^\perp u^\perp\|}{\|gu\|} \frac{\|u\|}{\|u^\perp\|}.$$

Furthermore, just as for (8.7),

$$(12.7) \quad -\log \text{VA}_r(gx, gx') \leq -\log \text{VA}_r(x, x') + Bn$$

for any  $x, x' \in \mathcal{F}(r, d)$  with  $F_1' \not\subset F_r$  and  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$ .

The proof of Proposition 12.1 is analogous to that of Proposition 8.1, replacing  $d(x, x')$ ,  $d(gx, gx')$  and  $d(x, E)$  with  $d(F_r, F_{r-1} + F_1')$ ,  $d(gF_r, g(F_{r-1} + F_1'))$  and  $d(F_r, E)$ , respectively, and substituting  $r$  for 1 in the subscript. The details follow, but the reader may choose to skip them and proceed directly to Section 12.3.

**Lemma 12.2.** *There exists  $\tilde{\kappa}_r = \tilde{\kappa}_r(\nu_{\mathcal{O}}) > 0$  and for each  $\delta > 0$  there exist  $\tilde{\theta}_r = \tilde{\theta}_r(\nu_{\mathcal{O}}, \delta) > 0$  and  $\tilde{N}_r = \tilde{N}_r(\nu_{\mathcal{O}}, \delta) \in \mathbb{N}$  such that for every  $n \geq \tilde{N}_r$  and  $x, x' \in \mathcal{F}(r, d)$  with  $F_1' \not\subset F_r$  there exists  $\tilde{\mathcal{E}}_r = \tilde{\mathcal{E}}_r(\nu_{\mathcal{O}}, \delta, n, x, x') \subset \text{supp } \nu_{\mathcal{O}}^{(n)}$  with  $\nu_{\mathcal{O}}^{(n)}(\tilde{\mathcal{E}}_r^c) < \delta$  and*

$$(12.8) \quad -\log \text{VA}_r(gx, gx') \leq \max\{-\log \text{VA}_r(x, x') - \tilde{\kappa}_r n, \tilde{\theta}_r\} \text{ for every } g \in \tilde{\mathcal{E}}_r.$$

*Proof.* Analogous to Lemma 8.2. Let  $\kappa_0 = \kappa_0(\nu_{\mathcal{O}}) > 0$ ,  $N_0 = N_0(\nu_{\mathcal{O}}, \delta) \in \mathbb{N}$ , and  $\mathcal{E}_0 = \mathcal{E}_0(\nu_{\mathcal{O}}, \delta, n, u^\perp) \subset \text{supp } \nu_{\mathcal{O}}^{(n)}$  and  $\tau_0 = \tau_0(\nu_{\mathcal{O}}, \delta) > 0$  be as in Proposition 5.5. Given  $x, x' \in \mathcal{F}(r, d)$  with  $F_1' \not\subset F_r$ , take  $u$  to be a non-zero vector that realizes the supremum in the definition (12.2). Write  $u = u^E + u^\perp$  with  $u^E \in E$  and  $u^\perp \in E^\perp$ . Take

$$(12.9) \quad \begin{aligned} \tilde{\kappa}_r &= \kappa_0/2, & \tilde{\theta}_r &= -\log(\tau_0/4), \\ \tilde{N}_r &> \max\{N_0, 4/\tilde{\kappa}_r\}, & \text{and } \tilde{\mathcal{E}}_r &= \mathcal{E}_0(\nu_{\mathcal{O}}, \delta, n, u^\perp). \end{aligned}$$

Let  $n \geq \tilde{N}_r$  and  $g \in \tilde{\mathcal{E}}_r$ ; then  $g \in \text{supp } \nu_{\mathcal{X}}^{(n)}$ . If  $\|g^\perp u^\perp\|/\|gu\| \geq \tau_0/2$  then the first inequality in (12.6) implies that

$$(12.10) \quad -\log \text{VA}_r(gx, gx') \leq -\log \frac{\tau_0}{2} \leq \tilde{\theta}_r.$$

If  $\|g^\perp u^\perp\|/\|gu\| < \tau_0/2$  then part (2) of Proposition 5.5 gives that

$$\frac{\|gu^\perp\|}{\|gu\|} < \frac{1}{2}, \text{ and so } \frac{\|gu^E\|}{\|gu\|} > \frac{1}{2}.$$

Substituting the latter inequality and  $\|u\| \geq \|u^E\|$  in (12.6), we find that

$$(12.11) \quad \text{VA}_r(gx, gx') \geq \frac{1}{2} \text{VA}_r(x, x') \frac{\|g^\perp u^\perp\|}{\|u^\perp\|} \frac{\|u^E\|}{\|gu^E\|}.$$

Thus, recalling the definition (5.3) and part (1) of Proposition 5.5,

$$\text{VA}_r(gx, gx') \geq \frac{1}{2} \text{VA}_r(x, x') \frac{\|Dg_{u^E}^\perp u^\perp\|}{\|u^\perp\|} \geq \frac{1}{2} \text{VA}_r(x, x') e^{\kappa_0 n}.$$

By the choices of  $\tilde{\kappa}_r$  and  $\tilde{N}_r$  in (12.9), this implies

$$(12.12) \quad \begin{aligned} -\log \text{VA}_r(gx, gx') &\leq -\log \text{VA}_r(x, x') + \log 2 - 2\tilde{\kappa}_r n \\ &\leq -\log \text{VA}_r(x, x') - \tilde{\kappa}_r n. \end{aligned}$$

The conclusion of the lemma is contained in (12.10) and (12.12).  $\square$

**12.2. Vertical projection function.** The *vertical projection function*  $\text{VP}_r$  is defined by

$$(12.13) \quad \text{VP}_r(x, x') = \text{VA}_r(x, x') d(F_r, F_{r-1} + F_1')^{\gamma_r}$$

where  $\gamma_r = \gamma_r(\nu_{\mathcal{X}})$  is a small positive constant chosen through the following result, which extends Proposition 8.3 to  $r > 1$ :

**Proposition 12.3.** *There exist  $\gamma_r = \gamma_r(\nu_{\mathcal{X}}) > 0$  and  $\kappa'_r = \kappa'_r(\nu_{\mathcal{X}}) > 0$  and for each  $\delta > 0$  there exists  $N_r = N_r(\nu_{\mathcal{X}}, \delta) \in \mathbb{N}$  such that for every  $n \geq N_r$  there exists  $\rho'_r = \rho'_r(\nu_{\mathcal{X}}, \delta, n) > 0$  such that for any  $x, x' \in E_r^{\mathcal{X}}(\rho'_r)$  with  $F_1' \not\subset F_r$  there exists  $\mathcal{E}'_r = \mathcal{E}'_r(\nu_{\mathcal{X}}, \delta, n, x, x') \subset \text{supp } \nu_{\mathcal{X}}^{(n)}$  with  $\nu_{\mathcal{X}}^{(n)}((\mathcal{E}'_r)^c) < \delta$  and*

$$(12.14) \quad -\log \text{VP}_r(gx, gx') \leq -\log \text{VP}_r(x, x') - \kappa'_r n \text{ for every } g \in \mathcal{E}'_r.$$

For the proof of Proposition 12.3 we need to extend Lemmas 8.4 and 8.5 to  $r > 1$ , which we do in the couple of statements that follow.

Let  $v \in F_r$  and  $v' \in F_{r-1} + F_1'$  be unit vectors orthogonal to  $F_{r-1}$  such that

$$(12.15) \quad \begin{aligned} F_r &= F_{r-1} + \mathbb{R}v, \\ F_{r-1} + F_1' &= F_{r-1} + \mathbb{R}v', \text{ and} \\ d(F_r, F_{r-1} + F_1') &= |\sin \angle(v, v')|. \end{aligned}$$

It is no restriction to take the angle between  $v$  and  $v'$  to be non-obtuse. Then the vector  $w = v' - v$  satisfies (compare Figure 5)

$$(12.16) \quad d(F_r, F_{r-1} + F_1') \leq \|w\| \text{ and } \angle(w, v) \geq \frac{\pi}{4}.$$

**Lemma 12.4.** *Given  $x, x' \in \mathcal{F}(r, d)$  with  $F'_1 \not\subset F_r$ , let  $w = v' - v$  be as in (12.16). Then*

$$(12.17) \quad d(F_r, E) < \frac{1}{4} \text{VA}_r(x, x') \text{ implies } \frac{\|w^\perp\|}{\|w\|} > \frac{1}{4} \text{VA}_r(x, x').$$

*Proof.* Analogous to Lemma 8.4. Let  $u \in F'_1 + F_r$  be any non-zero vector that realizes the supremum in the definition (12.2). Since  $F'_1 + F_r = \text{span}\{F_{r-1}, v, w\}$ , we may write  $u = u_0 + av + bw$  with  $u_0 \in F_{r-1}$  and  $a, b \in \mathbb{R}$ . Recall that  $v$  and  $w$  are orthogonal to  $F_{r-1}$ . Moreover, by (12.16) the angle between them is no less than  $\pi/4$ . This implies that  $\|u_0 + av\|$  and  $\|bw\|$  are both less than  $2\|u\|$ . Thus,

$$\begin{aligned} \text{VA}_r(x, x') &= \frac{\|u^\perp\|}{\|u\|} \leq \frac{\|(u_0 + av)^\perp\|}{\|u\|} + \frac{\|bw^\perp\|}{\|u\|} \\ &< 2 \frac{\|(u_0 + av)^\perp\|}{\|u_0 + av\|} + 2 \frac{\|w^\perp\|}{\|w\|} \leq 2d(F_r, E) + 2 \frac{\|w^\perp\|}{\|w\|}. \end{aligned}$$

Thus,  $d(F_r, E)$  and  $\|w^\perp\|/\|w\|$  cannot be both less than  $\text{VA}_r(x, x')/4$ .  $\square$

Take  $\tilde{\kappa}_r = \tilde{\kappa}_r(\nu_\infty) > 0$  and  $\tilde{\theta}_r = \tilde{\theta}_r(\nu_\infty, \delta) > 0$  to be as in Lemma 12.2.

**Lemma 12.5.** *There exists  $\hat{\kappa}_r = \hat{\kappa}_r(\nu_\infty) > 0$  and for each  $\delta > 0$  there exists  $\hat{N}_r = \hat{N}_r(\nu_\infty, \delta) \in \mathbb{N}$  such that for each  $n \geq \hat{N}_r$  there exists  $\hat{\rho}_r = \hat{\rho}_r(\nu_\infty, \delta, n) > 0$  such that for any  $x, x' \in E_r^\circ(\hat{\rho}_r)$  with  $F'_1 \not\subset F_r$  and  $-\log \text{VA}_r(x, x') \leq \tilde{\theta}_r + \tilde{\kappa}_r n$  there exists  $\hat{\mathcal{E}}_r = \hat{\mathcal{E}}_r(\nu_\infty, \delta, n, x, x') \subset \text{supp } \nu_\infty^{(n)}$  with  $\nu_\infty^{(n)}(\hat{\mathcal{E}}_r^c) < \delta$  and*

$$(12.18) \quad -\log d(gF_r, g(F_{r-1} + F'_1)) \leq -\log d(F_r, F_{r-1} + F'_1) - \log \text{VA}_r(x, x') - \hat{\kappa}_r n$$

for every  $g \in \hat{\mathcal{E}}_r$ .

*Proof.* Analogous to Lemma 8.5. Let  $\kappa_0 = \kappa_0(\nu_\infty) > 0$ ,  $N_0 = N_0(\nu_\infty, \delta) \in \mathbb{N}$ , and  $\mathcal{E}_0 = \mathcal{E}_0(\nu_\infty, \delta, n, v^\perp) \subset \text{supp } \nu_\infty^{(n)}$  be as in Proposition 5.5. Given  $x$  and  $x'$  with  $F'_1 \not\subset F_r$ , let  $w = v' - v$  be as in (12.16). Take

$$(12.19) \quad \begin{aligned} \hat{\kappa}_r &= \tilde{\kappa}_r/2, \quad \hat{N}_r = \max\{N_0, 5/\hat{\kappa}_r\}, \\ \hat{\rho}_r &< e^{-\tilde{\theta}_r - \tilde{\kappa}_r n}/10, \text{ and } \hat{\mathcal{E}}_r = \mathcal{E}_0(\nu_\infty, \delta, n, w^\perp). \end{aligned}$$

Let  $n \geq \hat{N}_r$  and  $g \in \hat{\mathcal{E}}_r \subset \text{supp } \nu_\infty^{(n)}$ . As observed in (12.16),

$$(12.20) \quad d(F_r, F_{r-1} + F'_1) \leq \|w\|.$$

Let us suppose that  $\|gv\| \geq \|gv'\|$ ; the case  $\|gv\| \leq \|gv'\|$  is analogous, reversing the roles of  $F_r$  and  $F_{r-1} + F'_1$ . Just as in (8.22),

$$(12.21) \quad -\log d(gF_r, g(F_{r-1} + F'_1)) \leq -\log d(F_r, F_{r-1} + F'_1) - \log \frac{\|\Pi_{gv} gw\|}{\|gv\|} \frac{\|v\|}{\|w\|}.$$

for any  $g \in G$ . By the condition on  $\hat{\rho}_r$  in (12.19), if  $d(F_r, E) \leq \hat{\rho}_r$  then

$$4d(F_r, E) < e^{-\tilde{\theta}_r - \tilde{\kappa}_r n} < \text{VA}_r(x, x'),$$

and then Lemma 12.4 gives that

$$(12.22) \quad \frac{\|w^\perp\|}{\|w\|} > \frac{1}{4} \text{VA}_r(x, x') > \frac{1}{4} e^{-\tilde{\theta}_r - \tilde{\kappa}_r n}.$$

Up to further reducing  $\hat{\rho}_r$ , we may also assume that

$$(12.23) \quad \frac{\|hz\|}{\|z\|} \geq \frac{1}{2} \frac{\|hz^E\|}{\|z^E\|} \text{ and } \|\Pi_{hz}hw\| \geq \frac{1}{2} \|\Pi_{hz^E}gw\|$$

for any non-zero  $z = z^E + z^\perp$  in  $E \oplus E^\perp$  with  $\|z^\perp\|/\|z\| \leq \hat{\rho}_r$  and any  $h \in \text{supp } \nu_\infty^{(n)}$ . The second part of (12.23) implies that

$$(12.24) \quad \|\Pi_{hz}hw\| \geq \frac{1}{2} \|\Pi_{hz^E}hw\| \geq \frac{1}{2} \|\Pi_Ehw\| = \frac{1}{2} \|(hw)^\perp\| = \frac{1}{2} \|h^\perp w^\perp\|.$$

Noting that  $\|v^\perp\|/\|v\| \leq d(F_r, E) \leq \hat{\rho}_r$ , take  $z = v$  and  $h = g$  in the previous two relations. Thus, substituting (12.23) and (12.24) in (12.21),

$$-\log d(gF_r, g(F_{r-1} + F'_1)) \leq -\log d(F_r, F_{r-1} + F'_1) + \log 4 - \log \frac{\|g^\perp w^\perp\|}{\|w\|} \frac{\|v^E\|}{\|gv^E\|}.$$

Then, using also (12.22) and (5.3),

$$\begin{aligned} & -\log d(gF_r, g(F_{r-1} + F'_1)) \\ & \leq -\log d(F_r, F_{r-1} + F'_1) - \log \text{VA}_r(x, x') + \log 16 - \log \frac{\|g^\perp w^\perp\|}{\|w^\perp\|} \frac{\|v^E\|}{\|gv^E\|} \\ & = -\log d(F_r, F_{r-1} + F'_1) - \log \text{VA}_r(x, x') + \log 16 - \log \frac{\|Dg_{v^E}^\perp w^\perp\|}{\|w^\perp\|}. \end{aligned}$$

By part (1) of Proposition 5.5 and the choice of  $\hat{\kappa}_r$  and  $\hat{N}_r$  in (12.19), it follows that

$$\begin{aligned} -\log d(gF_r, g(F_{r-1} + F'_1)) & \leq -\log d(F_r, F_{r-1} + F'_1) - \log \text{VA}_r(x, x') + 5 - 2\hat{\kappa}_r n \\ & \leq -\log d(F_r, F_{r-1} + F'_1) - \log \text{VA}_r(x, x') - \hat{\kappa}_r n, \end{aligned}$$

as claimed.  $\square$

*Proof of Proposition 12.3.* Analogous to Proposition 8.3. Take  $A = A(\nu_\infty)$  as in (8.2) and then define

$$(12.25) \quad \begin{aligned} \gamma_r &= \min \{1, \tilde{\kappa}_r/(2A)\}, \\ \kappa'_r &= \min \{\tilde{\kappa}_r/2, \gamma_r \tilde{\kappa}_r/2\}, \\ \rho'_r &= \min \{\hat{\rho}_r(\nu_\infty, \delta), \hat{\rho}_r(\nu_\infty, \delta/2)\}, \\ N_r &= \max \{\tilde{N}_r(\nu_\infty, \delta/2), \hat{N}_r(\nu_\infty, \delta/2), 2\tilde{\theta}_r/(\gamma \tilde{\kappa}_r)\}, \text{ and} \\ \mathcal{E}'_r &= \tilde{\mathcal{E}}_r(\nu_\infty, \delta/2, n, x, x') \cap \hat{\mathcal{E}}_r(\nu_\infty, \delta/2, n, x, x'). \end{aligned}$$

By construction,  $\nu_\infty^{(n)}((\mathcal{E}'_r)^c) < \delta$ . The definition (12.13) gives that

$$(12.26) \quad -\log \text{VP}_r(gx, gx') = -\log \text{VA}_r(gx, gx') - \gamma_r \log d(gF_r, g(F_{r-1} + F'_1)).$$

Consider  $x, x' \in E_r^\circ(\rho'_r)$  with  $F'_1 \not\subset F$ , and let  $n \geq N_r$  and  $g \in \mathcal{E}'_r$ . First, suppose that  $-\log \text{VA}_r(x, x') \geq \theta_r + \tilde{\kappa}_r n$ . Then, by Lemma 12.2,

$$(12.27) \quad -\log \text{VA}_r(gx, gx') \leq -\log \text{VA}_r(x, x') - \tilde{\kappa}_r n$$

Substituting (12.27) and (8.2) in (12.26) we find that

$$-\log \text{VP}_r(gx, gx') \leq -\log \text{VA}_r(x, x') - \tilde{\kappa}_r n - \gamma_r \log d(F_r, F_{r-1} + F'_1) + \gamma_r An.$$

By (12.26) and the choice of  $\gamma_r$  and  $\kappa'_r$  in (12.25), this yields

$$(12.28) \quad -\log \text{VP}_r(gx, gx') \leq -\log \text{VP}_r(x, x') - \frac{\tilde{\kappa}_r}{2}n \leq -\log \text{VP}_r(x, x') - \kappa'_r n.$$

Now assume that  $-\log \text{VA}_r(x, x') \leq \tilde{\theta}_r + \tilde{\kappa}_r n$ . In this case, Lemma 12.2 yields

$$(12.29) \quad -\log \text{VA}_r(gx, gx') \leq \tilde{\theta}_r$$

and Lemma 12.5 gives that

$$(12.30) \quad -\log d(gF_r, g(F_{r-1} + F'_1)) \leq -\log d(F_r, F_{r-1} + F'_1) - \log \text{VA}_r(x, x') - \hat{\kappa}_r n.$$

Substituting (12.29) and (12.30) in (12.26) we obtain

$$\begin{aligned} -\log \text{VP}_r(gx, gx') &\leq \tilde{\theta}_r - \gamma_r \log d(F_r, F_{r-1} + F'_1) - \gamma_r \log \text{VA}_r(x, x') - \gamma_r \hat{\kappa}_r n \\ &\leq -\log \text{VP}_r(x, x') + \tilde{\theta}_r + (1 - \gamma_r) \log \text{VA}_r(x, x') - \gamma_r \hat{\kappa}_r n. \end{aligned}$$

Since  $\text{VA}_r(x, x') \leq 1$ ,  $\gamma_r \leq 1$ , and  $n \geq N_r \geq 2\tilde{\theta}_r/(\gamma_r \hat{\kappa}_r)$ , this yields

$$(12.31) \quad \begin{aligned} -\log \text{VP}_r(gx, gx') &\leq -\log \text{VP}_r(x, x') + \tilde{\theta}_r - \gamma_r \hat{\kappa}_r n \\ &\leq -\log \text{VP}_r(x, x') - \frac{\gamma_r \hat{\kappa}_r}{2}n \leq -\log \text{VP}_r(x, x') - \kappa'_r n. \end{aligned}$$

The relations (12.28) and (12.31) contain the conclusion of Proposition 12.3.  $\square$

**12.3. The function  $-\log \psi_r$ .** At this point, the proof of Proposition 12.1 is analogous to that of Proposition 8.1. Take  $\kappa'_r > 0$ ,  $N_r \in \mathbb{N}$  and  $\rho'_r > 0$  as in Proposition 12.3. Let  $C'_r = C'_r(\nu_\infty) > 0$  be given by

$$(12.32) \quad C'_r = B + \gamma_r A.$$

Substituting (12.7) and (8.2) in the definition (12.13), we find that

$$(12.33) \quad \begin{aligned} -\log \text{VP}_r(gx, gx') &= -\log \text{VA}_r(gx, gx') - \gamma_r \log d(gF_r, g(F_{r-1} + F'_1)) \\ &\leq -\log \text{VA}_r(x, x') + Bn - \gamma_r \log d(F_r, F_{r-1} + F'_1) + \gamma_r An \\ &= -\log \text{VP}_r(x, x') + C'_r n. \end{aligned}$$

for any  $g \in \text{supp } \nu_\infty^{(n)}$  and  $x, x' \in \mathcal{F}(r, d)$  with  $F'_1 \not\subset F_r$ . Integrating (12.14) over  $\mathcal{E}'_r$  and (12.33) over the complement, and using the fact that  $\nu_\infty^{(n)}((\mathcal{E}'_r)^c) < \delta$ , we get that

$$\int_G -\log \text{VP}_r(gx, gx') d\nu_\infty^{(n)}(g) \leq -\log \text{VP}_r(x, x') - \kappa'_r n + C'_r \delta n.$$

for every  $n \geq N_r$  and  $x, x' \in E_r^\circ(\rho'_r)$  with  $F'_1 \not\subset F_r$ . This completes the proof of Proposition 12.1.

However, as mentioned before, for  $r > 1$  it is possible to have  $x$  and  $x'$  with  $F_r$  and  $F'_r$  arbitrarily close to the equator  $E$  without  $\text{VP}_r(x, x')$  getting close to zero. Here is a simple example (see also part (2) of Lemma 12.7):

**Example 12.6.** Denote by  $(x_1, x_2, x_3, x_4)$  the elements of  $\mathbb{R}^4$  and let  $E = \{x_4 = 0\}$ ,  $F_1 = \{x_2 = x_3 = x_4 = 0\}$ ,  $F_2 = \{x_3 = x_4 = 0\}$ ,  $F'_1 = \{x_2 = x_3 = 0, x_4 = \varepsilon x_1\}$  and  $F'_2 = \{x_3 = 0, x_4 = \varepsilon x_1\}$ . It is clear that  $F_2 \subset E$  and  $F'_2 \rightarrow E$  when  $\varepsilon \rightarrow 0$ . However,

- $\text{VA}_2(x, x') = d(F_2 + F'_1, E) = 1$ , since  $(0, 0, 0, 1) \in (F_2 + F'_1) \cap E^\perp$ ;
- $d(F_2, F_1 + F'_1) = 1$ , since  $(0, 1, 0, 0) \in F_2 \cap (F_1 + F'_1)^\perp$ .

It follows that  $\text{VP}_2(x, x')$  remains bounded from zero when  $\varepsilon \rightarrow 0$ .

Thus,  $-\log \text{VP}_r(x, x')$  cannot be used as a Margulis function for the equator  $E_r^\diamond$  in the flag space. To rectify this problem, we define the function  $\psi_r$  inductively in  $r$  as follows. Let  $x_-, x'_- \in \mathcal{F}_{r-1}(\mathbb{R}^d)$  be the *truncated flags* obtained by dropping the  $r$ -dimensional subspaces from  $x$  and  $x'$ , respectively. The assumption  $F'_1 \not\subset F_r$  implies that  $F'_1 \not\subset F_{r-1}$ , and so we may assume that  $\psi_{r-1}(x_-, x'_-)$  has already been defined. Then define

$$(12.34) \quad \psi_r(x, x') = \psi_{r-1}(x_-, x'_-)^{\beta_{r-1}} \text{VP}_r(x, x'),$$

where the exponent  $\beta_{r-1} = \beta_{r-1}(\nu_\infty)$  is a small constant to be chosen as follows.

It follows from this definition and (12.33) that

$$\begin{aligned} -\log \psi_j(gx, gx') + \log \psi_j(x, x') \\ \leq [-\log \psi_{j-1}(gx, gx') + \log \psi_{j-1}(x, x')] \beta_{j-1} + C'_j n \end{aligned}$$

for every  $j = 2, \dots, r$ . Thus, recalling also (8.35),

$$(12.35) \quad \begin{aligned} -\log \psi_r(gx, gx') \\ \leq -\log \psi_r(x, x') + [C'_1 \beta_1 \cdots \beta_{r-1} + \cdots + C'_{r-1} \beta_{r-1} + C'_r] n \\ \leq -\log \psi_r(x, x') + C''_r n, \end{aligned}$$

where  $C''_r = C''_r(\nu_\infty)$  is defined by

$$(12.36) \quad C''_r = C'_1 + \cdots + C'_{r-1} + C'_r.$$

Take the exponent  $\beta_{r-1}$  in (12.34) small enough that

$$(12.37) \quad \beta_{r-1} C''_{r-1} \leq \frac{1}{2} \kappa'_r$$

where  $\kappa'_r = \kappa'_r(\nu_\infty)$  is as in Proposition 12.1.

The function  $-\log \psi_r(x, x')$  thus defined does go to infinity when the flags  $x$  and  $x'$  approach the equator  $E_r^\diamond$ :

**Lemma 12.7.** *Given any  $R > 0$ , there exists  $\tilde{\rho}_r = \tilde{\rho}_r(\nu_\infty, R) > 0$  such that for any  $x, x' \in E_r^\diamond(\tilde{\rho}_r)$  with  $F'_1 \not\subset F_r$ ,*

- (1)  $-\log \psi_r(x, x') > R$  and
- (2)  $-\log \text{VP}_r(x, x') > R$  unless  $\text{VA}_{r-1}(x_-, x'_-) > \tilde{\rho}_r$ .

*Proof.* It is clear from the definitions (12.2) and (12.13) that  $\text{VP}_j(x, x') \leq 1$  for every  $1 \leq j \leq r$  and any  $x, x' \in \mathcal{F}(j, d)$  with  $F'_1 \not\subset F_j$ . So, the definition (12.34) implies that

$$(12.38) \quad -\log \psi_r(x, x') \geq -\log \psi_1(F_1, F'_1) \beta_1 \cdots \beta_{r-1}.$$

By definition, the  $\beta_j$  depend only on  $\nu_\infty$ . Then Lemma 8.6 gives that for any  $R > 0$  there exists  $\rho > 0$  depending only on  $\nu_\infty$  and  $R$  such that the right-hand side of (12.38) is greater than  $R$  for any  $F_1 \neq F'_1$  in  $E_1(\rho)$ . Since  $d(F_1, E) \leq d(F_r, E)$  and  $d(F'_1, E) \leq d(F'_r, E)$ , because  $F_1 \subset F_r$  and  $F'_1 \subset F'_r$ , we get that  $-\log \psi_r(x, x') > R$  for any  $x, x' \in E_r^\diamond(\rho)$  with  $F'_1 \not\subset F_r$ . This proves part (1).

To prove part (2), consider  $x, x' \in \mathcal{F}(r, d)$  with  $d(F_r, E) \leq \rho$ ,  $d(F'_r, E) \leq \rho$ , and

$$(12.39) \quad \text{VA}_{r-1}(x_-, x'_-) = d(F_{r-1} + F'_1, E) \leq \rho.$$

If  $d(F_r, F'_{r-1} + F'_1) \leq \sqrt{\rho}$ , then

$$(12.40) \quad \text{VP}_r(x, x') \leq d(F_r, F'_{r-1} + F'_1)^{\gamma_r} \leq \rho^{\gamma_r/2}.$$

Now suppose that  $d(F_r, F'_{r-1} + F'_1) \geq \sqrt{\rho}$ . Let  $v$  and  $v'$  be unit vectors orthogonal to  $F_{r-1}$  as in (12.15). Take  $u \in F'_1 + F_r$  realizing the supremum in (12.2), and write  $u = u_0 + av + bv'$  with  $u_0 \in F_{r-1}$  and  $a, b \in \mathbb{R}$ . Since  $|\sin \angle(v', F_r)| = d(F_r, F'_1 + F'_{r-1})$  is taken to be greater than  $\sqrt{\rho}$ ,

$$\|u_0 + av\| \leq \frac{2}{\sqrt{\rho}} \|u\| \text{ and } \|bv'\| \leq \frac{2}{\sqrt{\rho}} \|u\|.$$

Then,

$$\begin{aligned} \text{VA}_r(x, x') &= \frac{\|u^\perp\|}{\|u\|} \leq \frac{\|(u_0 + av)^\perp\|}{\|u\|} + \frac{\|(bv')^\perp\|}{\|u\|} \\ &\leq \frac{2}{\sqrt{\rho}} \frac{\|(u_0 + av)^\perp\|}{\|u_0 + av\|} + \frac{2}{\sqrt{\rho}} \frac{\|(v')^\perp\|}{\|v'\|}. \end{aligned}$$

Recalling also (12.39), we get that

$$(12.41) \quad \text{VP}_r(x, x') \leq \text{VA}_r(x, x') \leq \frac{2}{\sqrt{\rho}} d(F_r, E) + \frac{2}{\sqrt{\rho}} d(F'_1 + F'_{r-1}, E) \leq 4\sqrt{\rho}.$$

Since  $\gamma_r$  and  $\rho$  may be taken to be smaller than 1, both inequalities (12.40) and (12.41) imply that

$$-\log \text{VP}_r(x, x') \geq -\frac{\gamma_r}{2} \log \rho - \log 4$$

for any  $x, x'$  in  $E_r^\circ(\rho)$  with  $F'_1 \not\subset F_r$ . The right-hand side is larger than  $R$ , as long as  $\rho$  is chosen to be sufficiently small.  $\square$

For completeness, we include the following version of Propositions 12.1 and 12.3 for the function  $\psi_r$ , although it will not be needed in what follows (the related Proposition 13.2 will be used instead):

**Proposition 12.8.** *There exists  $\kappa_r'' = \kappa_r''(\nu_\infty) > 0$  such that for each  $\delta > 0$  and  $n \geq N_r$  there exists  $\rho_r'' = \rho_r''(\nu_\infty, \delta, n) > 0$  such that for any  $x, x' \in E_r^\circ(\rho_r'')$  with  $F'_1 \not\subset F_r$  there exists  $\mathcal{E}_r'' = \mathcal{E}_r''(\nu_\infty, \delta, n, x, x') \subset \text{supp } \nu_\infty^{(n)}$  with  $\nu_\infty^{(n)}((\mathcal{E}_r'')^c) < \delta$  and*

$$(12.42) \quad -\log \psi_r(gx, gx') \leq -\log \psi_r(x, x') - \kappa_r'' n \text{ for every } g \in \mathcal{E}_r''$$

and

$$(12.43) \quad \int_G -\log \psi_r(gx, gx') d\nu_\infty^{(n)}(g) \leq -\log \psi_r(x, x') - (\kappa_r'' - C_r'' \delta) n.$$

*Proof.* The case  $r = 1$  consists of the inequalities (8.15) and (8.1), respectively, with  $\kappa_1'' = \kappa_1'$ ,  $C_1'' = C_1'$ ,  $\rho_1'' = \rho_1'$ , and  $\mathcal{E}_1'' = \mathcal{E}_1'$ . Now suppose that  $r > 1$ . Recall that  $C_r'' = C_r''(\nu_\infty)$  was defined in (12.36). Define also

$$(12.44) \quad \kappa_r'' = \kappa_r'(\nu_\infty)/2, \quad \rho_r'' = \rho_r'(\nu_\infty, \delta, n) \text{ and } \mathcal{E}_r'' = \mathcal{E}_r'(\nu_\infty, \delta, n, x, x').$$

Consider  $n \geq N_r$ , and  $x, x' \in E_r^\circ(\rho_r'')$  with  $F'_1 \not\subset F_r$ . By (12.35),

$$(12.45) \quad -\log \psi_{r-1}(gx_-, gx'_-) \leq -\log \psi_{r-1}(x_-, x'_-) + C_{r-1}'' n$$

for every  $g \in \text{supp } \nu_\infty^{(n)}$ . By Proposition 12.3,

$$(12.46) \quad -\log \text{VP}_r(gx, gx') \leq -\log \text{VP}_r(x, x') - \kappa_r' n$$

for every  $g \in \mathcal{E}_r''$ . Substituting (12.45) and (12.46) in the definition (12.34), and recalling the choice of  $\beta_{r-1}$  in (12.37),

$$\begin{aligned} & -\log \psi_r(gx, gx') - \log \psi_{r-1}(gx_-, gx'_-) \beta_{r-1} - \log \text{VP}_r(gx, gx') \\ & \leq -\log \psi_{r-1}(x_-, x'_-) \beta_{r-1} + C_{r-1}'' n \beta_{r-1} - \log \text{VP}_r(x, x') - \kappa_r' n \\ & \leq -\log \psi_r(x, x') - (\kappa_r'/2)n \end{aligned}$$

for every  $g \in \mathcal{E}_r''$ . This gives claim (12.42). Moreover, integrating this inequality on  $\mathcal{E}_r''$  and (12.35) on the complement, we obtain

$$\int_G -\log \psi_r(gx, gx') d\nu_{\mathcal{O}}^{(n)}(g) \leq -\log \psi_r(x, x') - (\kappa_r'/2 - C_r'' \delta)n,$$

which gives claim (12.43).  $\square$

### 13. STABILIZATION AND CUT-OFF

Next we present the analogues for  $r > 1$  of the two constructions in Section 9. Both are fairly straightforward extensions of the case  $r = 1$ , but it turns out that for  $r > 1$  they do not suffice to deal with the questions discussed at the beginning of Section 9. This difficulty will be handled later, in Section 15. Another issue is that the function  $\psi_r$  we construct in the following is not symmetric when  $r > 1$ . Thus we will use instead the function  $\hat{\psi}_r$  defined by

$$\hat{\psi}_r(x, x') = \max\{\psi_r(x, x'), \psi_r(x', x)\}.$$

Let  $\kappa_r'' = \kappa_r''(\nu_{\mathcal{O}}) > 0$ ,  $C_r'' = C_r''(\nu_{\mathcal{O}}) > 0$  and  $N_r = N_r(\nu_{\mathcal{O}}, \delta) \in \mathbb{N}$  be as in Proposition 12.8. Keep in mind that  $\varepsilon_r < \rho_r'' = \rho_r' \leq \hat{\rho}_r$ ,  $C_r' = B + \gamma_r A$ , and  $C_r'' = C_1' + \dots + C_r'$ , according to (12.25), (12.32), (12.36), (12.44), and (13.17).

**13.1. Stabilization.** Let  $\gamma_r = \gamma_r(\nu_{\mathcal{O}}) > 0$  and  $B = B(\nu_{\mathcal{O}}) > 0$  be as in (12.25) and (8.3), respectively. For each  $\omega_r > 0$  and  $n \in \mathbb{N}$ , define the *stabilized vertical angle*

$$(13.1) \quad \text{SVA}_r(x, x'; \omega_r) = \max\{\text{VA}_r(x, x'), \omega_r e^{-Bn}\}$$

and the *stabilized vertical projection* by

$$(13.2) \quad \text{SVP}_r(x, x'; \omega_r) = \text{SVA}_r(x, x'; \omega_r) d(F_r, F_{r-1} + F_1')^{\gamma_r}$$

for every  $x, x'$  in  $E_r^\diamond(\varepsilon_r)$  with  $F_1' \not\subset F_r$ .

Given  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$ , let  $\vec{\omega}_- = (\omega_1, \dots, \omega_{r-1})$  and define the *stabilized function*  $\psi_r$  by

$$(13.3) \quad \psi_r(x, x'; \vec{\omega}) = \psi_{r-1}(x_-, x'_-; \vec{\omega}_-)^{\beta_{r-1}} \text{SVP}_r(x, x'; \omega_r).$$

The following extension of Lemma 9.1 asserts that the estimate in (12.35) remains valid for these stabilized functions:

**Lemma 13.1.** *For every  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$ , any  $x, x' \in E_r^\diamond(\rho_r'')$  with  $F_1' \not\subset F_r$ , and any  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$ ,*

$$(13.4) \quad -\log \psi_r(gx, gx'; \vec{\omega}) \leq -\log \psi_r(x, x'; \vec{\omega}) + C_r'' n.$$



*Proof.* The case  $r = 1$  was done in Lemma 9.1, so let us consider  $r > 1$ . By induction,

$$(13.5) \quad -\log \psi_r(gx_-, gx'_-; \vec{\omega}_-) \leq -\log \psi_r(x_-, x'_-; \vec{\omega}_-) + C''_{r-1}n.$$

We claim that

$$(13.6) \quad -\log \text{SVA}_r(gx, gx'; \omega_r) \leq -\log \text{SVA}_r(x, x'; \omega_r) + Bn.$$

and

$$(13.7) \quad -\log \text{SVP}_r(gx, gx'; \omega_r) \leq -\log \text{SVP}_r(x, x'; \omega_r) + C'_r n.$$

The inequality (13.4) follows directly from combining (13.5) and (13.7), and recalling the definition of  $C''_r$  in (12.36):

$$(13.8) \quad \begin{aligned} -\log \psi_r(gx, gx'; \vec{\omega}) &= -\log \psi_r(gx_-, gx'_-; \vec{\omega}_-) \beta_{r-1} \\ &\quad - \log \text{SVP}_r(gx, gx'; \omega_r) \\ &\leq -\log \psi_r(x_-, x'_-; \vec{\omega}_-) \beta_{r-1} + C''_{r-1} n \beta_{r-1} \\ &\quad - \log \text{SVP}_r(x, x'; \omega_r) + C'_{r-1} n \\ &\leq -\log \psi_r(x, x'; \vec{\omega}) + C''_r n. \end{aligned}$$

Recall also that  $\beta_{r-1} \leq 1$ .

We split the proof of (13.6) into two cases. Suppose first that  $\text{VA}_r(x, x') < \omega_r$ . Then, by the definition (13.1),

$$\begin{aligned} -\log \text{SVA}_r(x, x'; \omega_r) &\geq -\log \omega_r \text{ and} \\ -\log \text{SVA}_r(gx, gx'; \omega_r) &\leq -\log \omega_r + Bn \leq -\log \text{SVA}_r(x, x'; \omega_r) + Bn \end{aligned}$$

as claimed. Now suppose that  $\text{VA}_r(x, x') \geq \omega_r$ . Then, again by the definition (13.1),

$$\begin{aligned} -\log \text{SVA}_r(x, x'; \omega_r) &= -\log \text{VA}_r(x, x') \text{ and} \\ -\log \text{SVA}_r(gx, gx'; \omega_r) &\leq -\log \text{VA}_r(gx, gx'). \end{aligned}$$

Together with (12.7), this yields

$$-\log \text{SVA}_r(gx, gx'; \omega_r) \leq -\log \text{SVA}_r(x, x'; \omega_r) + Bn,$$

which completes the proof of (13.6).

Finally, substituting (13.6) and (8.2) in the definition (12.13), and recalling the definition of  $C'_r$  in (12.32),

$$(13.9) \quad \begin{aligned} -\log \text{SVP}_r(gx, gx'; \omega_r) &= -\log \text{SVA}_r(gx, gx'; \omega_r) \\ &\quad - \gamma_r \log d(gF_r, g(F_{r-1} + F'_1)) \\ &\leq -\log \text{SVA}_r(x, x'; \omega_r) + Bn \\ &\quad - \gamma_r \log d(gF_r, g(F_{r-1} + F'_1)) - \gamma_r An \\ &= -\log \text{SVP}_r(x, x'; \omega_r) + (B + \gamma_r A)n \end{aligned}$$

This proves (13.7), and thus completes the proof of the lemma.  $\square$

We say that  $(x, x')$  is in the *stabilization region* if  $\text{VA}_r(x, x') < \omega_r$ . As we have seen in (12.7),

$$\text{VA}_r(gx, gx') \geq \text{VA}_r(x, x') e^{-Bn}$$

for any  $x, x' \in \mathcal{F}(r, d)$  with  $F_1' \not\subset F_r$  and  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$ . So, if  $(x, x')$  is outside the stabilization region then

$$(13.10) \quad \begin{aligned} \text{SVA}_r(x, x'; \omega_r) &= \text{VA}_r(x, x') \text{ and } \text{SVA}_r(gx, gx'; \omega_r) = \text{VA}_r(gx, gx') \\ \text{SVP}_r(x, x'; \omega_r) &= \text{VP}_r(x, x') \text{ and } \text{SVP}_r(gx, gx'; \omega_r) = \text{VP}_r(gx, gx') \end{aligned}$$

for any  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$  and  $n \in \mathbb{N}$ .

The following analogue of Proposition 12.8 for stabilized vertical angles and the stabilized vertical projections extends Proposition 9.2 to  $r > 1$ :

**Proposition 13.2.** *For every  $\delta > 0$ ,  $n \geq N_r$ ,  $x, x'$  in  $E_r^\circ(\rho_r'')$  with  $F_1' \not\subset F_r$ , and  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$  with  $\text{VA}_r(x, x') \geq \omega_r$ ,*

$$(13.11) \quad -\log \psi_r(gx, gx'; \vec{\omega}) \leq -\log \psi_r(x, x'; \vec{\omega}) - \kappa_r'' n \text{ for every } g \in \mathcal{E}_r'',$$

and

$$(13.12) \quad \int_G -\log \psi_r(gx, gx'; \vec{\omega}) d\nu_{\mathcal{O}}^{(n)}(g) \leq -\log \psi_r(x, x'; \vec{\omega}) - (\kappa_r'' - C_r'' \delta) n$$

for any

*Proof.* The case  $r = 1$  is given by Proposition 9.2 with  $\kappa_1'' = \kappa_1'$ ,  $C_1'' = C_1'$ ,  $\rho_1'' = \rho_1'$ , and  $\mathcal{E}_1'' = \mathcal{E}_1'$ . Now let us consider  $r > 1$ . Let  $\delta > 0$ ,  $n \geq N_r$ ,  $x, x' \in E_r^\circ(\rho_r'')$ , and  $\vec{\omega} \in \mathbb{R}_+^d$  be as in the statement. By (13.3) and (13.10),

$$(13.13) \quad \begin{aligned} \psi_r(x, x'; \vec{\omega}) &= \psi_{r-1}(x_-, x'_-; \vec{\omega}_-)^{\beta_{r-1}} \text{VP}_r(x, x') \text{ and} \\ \psi_r(gx, gx'; \vec{\omega}) &= \psi_{r-1}(gx_-, gx'_-; \vec{\omega}_-)^{\beta_{r-1}} \text{VP}_r(gx, gx'). \end{aligned}$$

By (13.4),

$$(13.14) \quad -\log \psi_{r-1}(gx_-, gx'_-; \vec{\omega}_-) \leq -\log \psi_{r-1}(x_-, x'_-; \vec{\omega}_-) + C_{r-1}'' n$$

for every  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$ . By Proposition 12.3,

$$(13.15) \quad -\log \text{VP}_r(gx, gx') \leq -\log \text{VP}_r(x, x') - \kappa_r' n$$

for every  $g \in \mathcal{E}_r''$ . Substituting (13.14) and (13.15) in the second part of (13.13), and recalling the choice of  $\beta_{r-1}$  in (12.37),

$$\begin{aligned} -\log \psi_r(gx, gx'; \vec{\omega}) &= -\log \psi_{r-1}(gx_-, gx'_-; \vec{\omega}_-) \beta_{r-1} - \log \text{VP}_r(gx, gx') \\ &\leq -\log \psi_{r-1}(x_-, x'_-; \vec{\omega}_-) \beta_{r-1} + C_{r-1}'' n \beta_{r-1} \\ &\quad - \log \text{VP}_r(x, x') - \kappa_r' n \\ &\leq -\log \psi_r(x, x'; \vec{\omega}) - (\kappa_r'/2) n = -\log \psi_r(x, x'; \vec{\omega}) - \kappa_r'' n \end{aligned}$$

for every  $g \in \mathcal{E}_r''$ . This gives claim (13.11). Moreover, integrating (13.11) on  $\mathcal{E}_r''$  and (13.4) on the complement, we obtain that

$$\int_G -\log \psi_r(gx, gx'; \vec{\omega}) d\nu_{\mathcal{O}}^{(n)}(g) \leq -\log \psi_r(x, x'; \vec{\omega}) - (\kappa_r'' - C_r'' \delta) n,$$

as claimed in (13.12).  $\square$

**13.2. Cutoff.** Recall also that the constant  $\varepsilon_2 > 0$  was chosen at the end of the initial step of the induction, in Section 11.3. Recall also that it may be taken to be as small as we want.

For any  $r > 1$ , assume that  $\varepsilon_r = \varepsilon_r(\nu_{\mathcal{O}}, \delta, n) > 0$  has been chosen, satisfying

$$(13.16) \quad \eta_{\mathcal{O},r}(E_r(\varepsilon_r) \setminus E) < \frac{1}{10} \eta_{\mathcal{O},r}(E).$$

and

$$(13.17) \quad \varepsilon_r < \min\{\rho_0, \rho_r''\},$$

where  $\rho_0 = \rho_0(\nu_{\mathcal{O}}, n) > 0$  is as in Corollaries 5.6 and 5.7 and  $\rho_r'' = \rho_r''(\nu_{\mathcal{O}}, \delta, n) > 0$  is as in Proposition 13.2.

Using Remark 7.3 twice, we find  $\varepsilon_r' = \varepsilon_r'(\nu_{\mathcal{O}}, \delta, n) > 0$  and  $\tilde{\varepsilon}_r = \tilde{\varepsilon}_r(\nu_{\mathcal{O}}, \delta, n) > 0$  with  $0 < \varepsilon_r' < \tilde{\varepsilon}_r < \varepsilon_r$ , and a compact neighborhood  $\mathcal{W}_r = \mathcal{W}_r(\nu_{\mathcal{O}}, \delta, n)$  of  $\text{supp } \nu_{\mathcal{O}}^{(n)}$  such that

$$(13.18) \quad gF_r \in E_r(\varepsilon_r/2) \text{ for every } F_r \in E_r(2\tilde{\varepsilon}_r) \text{ and } g \in \mathcal{W}_r \text{ and}$$

$$(13.19) \quad g^{-1}F_r \in E_r(\tilde{\varepsilon}_r/2) \text{ for every } F_r \in E_r(2\varepsilon_r') \text{ and } g \in \mathcal{W}_r.$$

Let  $\kappa_0 = \kappa_0(\nu_{\mathcal{O}}) > 0$  be as in Proposition 5.5, and define  $\varepsilon_r'' = \varepsilon_r''(\nu_{\mathcal{O}}, \delta, n)$  by

$$(13.20) \quad \varepsilon_r'' = 3\varepsilon_r' e^{-\kappa_0 n/2}.$$

Taking  $\rho = \varepsilon_r''$  in Corollary 5.7, and keeping in mind that  $\varepsilon_r < \rho_0$ , we get that there are  $\tilde{k}_r = \tilde{k}_r(\nu_{\mathcal{O}}, \delta, n) \in \mathbb{N}$  and  $\mathcal{D}_k(F_r) = \mathcal{D}_k(\nu_{\mathcal{O}}, \delta, n, F_r) \subset \text{supp } \nu_k^{(n)}$  such that  $\nu_k^{(n)}(\mathcal{D}_k(F_r)^c) < \delta$  and

$$d(gF_r, E) > e^{\kappa_0 n/2} d(F_r, E) > e^{\kappa_0 n/2} \varepsilon_r'' > 2\varepsilon_r'$$

for any  $g \in \mathcal{D}_k(F_r)$ ,  $F_r \in E(\varepsilon_r, \varepsilon_r'')$  and  $k \geq \tilde{k}_r$ . In other words, for  $k \geq \tilde{k}_r$ ,

$$(13.21) \quad F_r \in E(\varepsilon_r, \varepsilon_r'') \Rightarrow gF_r \notin E(2\varepsilon_r') \text{ for every } g \in \mathcal{D}_k(F_r).$$

Increasing  $\tilde{k}_r$  if necessary, we may suppose that  $\text{supp } \nu_k^{(n)} \subset \mathcal{W}_r$  for every  $k \geq \tilde{k}_r$ . Then (13.18) and (13.19) imply

$$(13.22) \quad E_r(2\tilde{\varepsilon}_r) \subset \mathcal{X}_{\nu_k^{(n)}} E_r(\varepsilon_r) \text{ and } E_r(2\varepsilon_r') \subset \mathcal{X}_{\nu_k^{(n)}}^{\#} E_r(\varepsilon_r)$$

$$(13.23) \quad F_r \notin E_r(\tilde{\varepsilon}_r/2) \Rightarrow gF_r \notin E_r(2\varepsilon_r') \text{ for every } g \in \text{supp } \nu_k^{(n)}.$$

We say that  $x, x' \in \mathcal{F}(r, d)$  are *in general position* if  $F_1' \not\subset F_r$  and  $F_1 \not\subset F_r'$ . For any  $x, x' \in E_r^{\circ}(\varepsilon_r)$  in general position and  $\vec{\omega} \in \mathbb{R}_+^d$ , define

$$(13.24) \quad \Psi_r(x, x'; \vec{\omega}) = \begin{cases} \log \left( \Omega_r + \hat{\psi}_r(x, x'; \vec{\omega})^{-1} \right) & \text{if } x \in E_r^{\circ}(2\varepsilon_r'') \text{ or } x' \in E_r^{\circ}(2\varepsilon_r'') \\ \log \Omega_r & \text{otherwise,} \end{cases}$$

where  $\Omega_r = \Omega_r(\nu_{\mathcal{O}}, \delta, n) > 1$  is a large constant to be chosen in Proposition 15.1, and

$$(13.25) \quad \hat{\psi}(x, x'; \vec{\omega}) = \max\{\psi(x, x'; \vec{\omega}), \psi(x', x; \vec{\omega})\}.$$

It is clear from the definition that  $\Psi_r(\cdot, \cdot; \vec{\omega})$  is a symmetric function:

$$\Psi_r(x, x'; \vec{\omega}) = \Psi_r(x', x; \vec{\omega}) \text{ for all } x, x' \in E_r^{\circ}(\varepsilon_r).$$

The set  $E_r^{\circ}(2\varepsilon_r'')^c \times E_r^{\circ}(2\varepsilon_r'')^c$  is the *cut-off region* at the stage  $r > 1$ . Compare Figure 7.

**Proposition 13.3.** *There exists  $\kappa_r''' = \kappa_r'''(\nu_{\mathcal{O}}) > 0$  such that for each  $\delta > 0$  and  $n \geq N_r$  there exists  $\varepsilon_r'' = \varepsilon_r''(\nu_{\mathcal{O}}, \delta, n) > 0$  such that given any  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$ :*

(i) *For any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position with  $\Psi_r(x, x'; \vec{\omega}) > \log \Omega_r$ ,*

$$\int_G \Psi_r(gx, gx'; \vec{\omega}) d\nu_{\mathcal{O}}^{(n)}(g) \leq \Psi_r(x, x'; \vec{\omega}) + C_r'' n.$$

(ii) *For any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position with  $\Psi_r(x, x'; \vec{\omega}) > \log \Omega_r$ ,  $\text{VA}_r(x, x') \geq \omega_r$  and  $\text{VA}_r(x', x) \geq \omega_r$ ,*

$$\int_G \Psi_r(gx, gx'; \vec{\omega}) d\nu_{\mathcal{O}}^{(n)}(g) \leq \Psi_r(x, x'; \vec{\omega}) + C_r'' \delta n.$$

(iii) *For any  $x, x' \in E_r^\circ(\varepsilon_r''')$  in general position satisfying  $\text{VA}_r(x, x') \geq \omega_r$  and  $\text{VA}_r(x', x) \geq \omega_r$ ,*

$$\int_G \Psi_r(gx, gx'; \vec{\omega}) d\nu_{\mathcal{O}}^{(n)}(g) \leq \Psi_r(x, x'; \vec{\omega}) - (\kappa_r''' - C_r'' \delta)n.$$

*Proof.* Define  $\kappa_r''' = \kappa_r''/2$ . Let  $n \geq N_r$ . Part (i) of the proposition is a consequence of the following lemma:

**Lemma 13.4.** *If  $x, x' \in \mathcal{F}(r, d)$  are in general position and  $\Psi_r(x, x'; \vec{\omega}) > \log \Omega_r$  then*

$$(13.26) \quad \Psi_r(gx, gx'; \vec{\omega}) \leq \Psi_r(x, x'; \vec{\omega}) + C_r'' n \text{ for any } g \in \text{supp } \nu_{\mathcal{O}}^{(n)}.$$

*Proof.* It follows immediately from (13.4) that

$$(13.27) \quad -\log \hat{\psi}_r(gx, gx'; \vec{\omega}) \leq -\log \hat{\psi}_r(x, x'; \vec{\omega}) + C_r'' n$$

for any  $g \in \text{supp } \nu_{\mathcal{O}}^{(n)}$ . Then, using (9.1),

$$\begin{aligned} \Psi_r(gx, gx'; \vec{\omega}) &\leq \log \left( \Omega_r + \hat{\psi}_r(gx, gx'; \vec{\omega})^{-1} \right) \leq \log \left( \Omega_r + e^{C_r'' n} \hat{\psi}_r(x, x'; \vec{\omega})^{-1} \right) \\ &\leq \log \left( \Omega_r + \hat{\psi}_r(x, x'; \vec{\omega})^{-1} \right) + C_r'' n = \Psi_r(x, x'; \vec{\omega}) + C_r'' n. \end{aligned}$$

This proves the claim.  $\square$

To prove part (ii) we use Proposition 13.2: given any  $\vec{\omega} \in \mathbb{R}_+^d$  and  $x, x' \in E_r^\circ(\varepsilon_r)$  with  $F_1' \nmid F_r$ ,

$$(13.28) \quad -\log \psi_r(gx, gx'; \vec{\omega}) \leq -\log \psi_r(x, x'; \vec{\omega}) - \kappa_r'' n \leq -\log \psi_r(x, x'; \vec{\omega})$$

for every  $g \in \mathcal{E}_r''$ . This remains true if we exchange the roles of  $x$  and  $x'$ , of course.

Thus,  $-\log \hat{\psi}_r(gx, gx'; \vec{\omega}) \leq -\log \hat{\psi}_r(x, x'; \vec{\omega})$ , and so

$$(13.29) \quad \begin{aligned} \Psi_r(gx, gx'; \vec{\omega}) &\leq \log \left( \Omega_r + \hat{\psi}_r(gx, gx'; \vec{\omega})^{-1} \right) \\ &\leq \log \left( \Omega_r + \hat{\psi}_r(x, x'; \vec{\omega})^{-1} \right) = \Psi_r(x, x'; \vec{\omega}) \end{aligned}$$

for every  $g \in \mathcal{E}_r''$ . Integrating (13.29) over  $\mathcal{E}_r''$  and (13.26) over the complement, we obtain the estimate in part (ii).

Now take  $c = e^{-2\kappa_r''' n}$  in the relation (9.3). By Lemma 12.7, there exists  $\varepsilon_r''' > 0$  depending only on  $\nu_{\mathcal{O}}$ ,  $\delta$  and  $n$  (through  $c$  and  $\Omega_r$ ) such that

$$(13.30) \quad -\log \hat{\psi}_r(x, x'; \vec{\omega}) = -\log \hat{\psi}_r(x, x') \geq \log (\Omega_r / \sqrt{c})$$

for any  $x, x' \in E_r(\varepsilon_r''')$  in general position. Then, using (13.30), (13.28) and (9.3),

$$\begin{aligned}
(13.31) \quad \Psi_r(gx, gx'; \vec{\omega}) &\leq \log \left( \Omega_r + \hat{\psi}_r(gx, gx'; \vec{\omega})^{-1} \right) \\
&\leq \log \left( \Omega_r + e^{-2\kappa_r'' n} \hat{\psi}_r(x, x'; \vec{\omega})^{-1} \right) \\
&\leq \log \left( \Omega_r + \hat{\psi}_r(x, x'; \vec{\omega})^{-1} \right) - \kappa_r'' n = \Psi_r(x, x'; \vec{\omega}) - \kappa_r'' n
\end{aligned}$$

for every  $g \in \mathcal{E}_r''$ . Integrating (13.31) over  $\mathcal{E}_r''$  and (13.26) over the complement, we obtain the estimate in part (iii) of the proposition.  $\square$

#### 14. TURNING THE PERTURBATION ON

We show that the conclusions of Proposition 13.3 hold for  $\nu_k^{(n)}$  instead of  $\nu_{\mathcal{O}}^{(n)}$ , as long as  $k \in \mathbb{N}$  is sufficiently large. The arguments are close to those in Section 10, but we have to deal with the fact that  $\psi_r$  and  $\Psi_r$  are not entirely straightforward generalizations of  $\psi_1$  and  $\Psi_1$ .

**Proposition 14.1.** *Given  $\delta > 0$ ,  $n \geq N_r$ , and  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$  there is  $k_r = k_r(\nu_{\mathcal{O}}, \delta, n, \omega_r) \in \mathbb{N}$  such that the following holds for every  $k \geq k_r$ :*

(i) *For any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position with  $\Psi_r(x, x'; \vec{\omega}) > \log \Omega_r$ ,*

$$\int_G \Psi_r(gx, gx'; \vec{\omega}) d\nu_k^{(n)}(g) \leq \Psi_r(x, x'; \vec{\omega}) + C_r'' n.$$

(ii) *For any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position with  $\Psi_r(x, x'; \vec{\omega}) > \log \Omega_r$ ,  $\text{VA}_r(x, x') \geq \omega_r$ , and  $\text{VA}_r(x', x) \geq \omega_r$ ,*

$$\int_G \Psi_r(gx, gx'; \vec{\omega}) d\nu_k^{(n)}(g) \leq \Psi_r(x, x'; \vec{\omega}) + C_r'' \delta n.$$

(iii) *For any  $x, x' \in E_r^\circ(\varepsilon_r''')$  in general position satisfying  $\text{VA}_r(x, x') \geq \omega_r$  and  $\text{VA}_r(x', x) \geq \omega_r$ ,*

$$\int_G \Psi_r(gx, gx'; \vec{\omega}) d\nu_k^{(n)}(g) \leq \Psi_r(x, x'; \vec{\omega}) - (\kappa_r'' - C_r'' \delta) n.$$

Keep in mind that we have chosen  $\varepsilon_r < \rho_r'' = \rho_r' \leq \hat{\rho}_r$  and  $C_r' = B + \gamma_r A$ . Recall also that  $\tilde{k}_r = \tilde{k}_r(\nu_{\mathcal{O}}, \delta, n) \in \mathbb{N}$  was chosen so that the relations (13.21) through (13.23) hold for every  $k \geq \tilde{k}_r$ .

*Proof.* We are going to extend to large  $k \in \mathbb{N}$  several estimates in the proof of Proposition 13.3. This will require a number of conditions on  $k$ , depending on  $\nu_{\mathcal{O}}$ ,  $\delta$ ,  $n$  and  $\omega$ , that we state along the way. We begin with the following extension of Lemma 13.4:

**Lemma 14.2.** *Given any  $n \geq N_r$  and  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$ , there exists  $\hat{k}_r = \hat{k}_r(\nu_{\mathcal{O}}, n, \omega_r)$  such that if  $x, x' \in E_r^\circ(\varepsilon_r)$  are in general position and  $\vec{\omega} \in \mathbb{R}_+^d$  is such that  $\Psi_r(x, x'; \vec{\omega}) > \log \Omega_r$  then*

$$(14.1) \quad \Psi_r(gx, gx'; \vec{\omega}) \leq \Psi_r(x, x'; \vec{\omega}) + C_r'' n$$

for any  $g \in \text{supp } \nu_k^{(n)}$  and  $k \geq \hat{k}_r$ .

*Proof.* Consider  $\delta > 0$ ,  $n \geq N_r$ , and  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$ . We begin by claiming that there exists  $\hat{k}_r = \hat{k}_r(\nu_\infty, n, \vec{\omega}) \in \mathbb{N}$  such that

$$(14.2) \quad -\log \text{SVA}_r(gx, gx'; \omega_r) \leq -\log \text{SVA}_r(x, x'; \omega_r) + Bn$$

for any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position,  $g \in \text{supp } \nu_k^{(n)}$ , and  $k \geq \hat{k}_r$ . This can be seen as follows. If  $\text{VA}_r(x, x') < \omega_r$  then, by the definition (13.1),

$$\begin{aligned} -\log \text{SVA}_r(x, x'; \omega_r) &\geq -\log \omega_r \text{ and} \\ -\log \text{SVA}_r(gx, gx'; \omega_r) &\leq -\log \omega_r + Bn \leq -\log \text{SVA}_r(x, x'; \omega_r) + Bn, \end{aligned}$$

as claimed. Now suppose that  $\text{VA}_r(x, x') \geq \omega_r$ . The relation (12.7) does not apply here. Instead, from (12.5) and (13.1) we get that

$$(14.3) \quad -\log \text{SVA}_r(gx, gx'; \omega_r) \leq -\log \text{SVA}_r(x, x'; \omega_r) - \log \frac{\|(gu)^\perp\| \|u\|}{\|gu\| \|u^\perp\|}$$

for every  $g \in \text{supp } \nu_k^{(n)}$ , with  $u = u(x, x')$  realizing the supremum in (12.2). The assumption  $\text{VA}_r(x, x') \geq \omega_r$  means that  $\|u^\perp\| \geq \omega_r \|u\|$ . Since  $\text{supp } \nu_k$  converges to  $\text{supp } \nu_\infty$  in the Hausdorff topology, we may find  $r_k = r_k(\nu_\infty, n) \rightarrow 0$  such that every  $g \in \text{supp } \nu_k^{(n)}$  is in the  $r_k$ -neighborhood of some  $f \in \text{supp } \nu_\infty^{(n)}$ . Then

$$\|(gu)^\perp - (fu)^\perp\| \leq \|gu - fu\| \leq r_k \|u\| \leq \frac{r_k}{\omega_r} \|u^\perp\|.$$

Then we may take  $\hat{k}_r = \hat{k}_r(\nu_\infty, n, \omega_r) \in \mathbb{N}$  large enough so that this inequality implies

$$\begin{aligned} -\log \frac{\|(gu)^\perp\| \|u\|}{\|gu\| \|u^\perp\|} &\leq -\log \frac{\|(fu)^\perp\| \|u\|}{\|fu\| \|u^\perp\|} + \log 2 \\ &= -\log \frac{\|f^\perp u^\perp\| \|u\|}{\|fu\| \|u^\perp\|} + \log 2 \\ &\leq \log \|(f^\perp)^{-1}\| + \log \|f\| + \log 2 \leq Bn \end{aligned}$$

for every  $g \in \text{supp } \nu_k^{(n)}$  and  $k \geq \hat{k}_r$ . This proves (14.2).

Substituting (14.2) and (8.2) in the definition (13.2), and recalling the definition of  $C'_r$  in (12.32), we find that

$$(14.4) \quad \begin{aligned} -\log \text{SVP}_r(gx, gx'; \omega_r) &= -\log \text{SVA}_r(gx, gx'; \omega_r) \\ &\quad - \gamma_r \log d(gF_r, g(F_{r-1} + F'_1)) \\ &\leq -\log \text{SVA}_1(x, x'; \omega_r) + Bn \\ &\quad - \gamma_r \log d(F_r, F_{r-1} + F'_1) + \gamma_r An \\ &\leq -\log \text{SVP}_r(x, x'; \omega_r) + C'_r n \end{aligned}$$

for any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position,  $g \in \text{supp } \nu_k^{(n)}$ , and  $k \geq \hat{k}_r$ .

Next we claim that

$$(14.5) \quad -\log \psi_r(gx, gx'; \vec{\omega}) \leq -\log \psi_r(x, x'; \vec{\omega}) + C''_r n$$

for any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position,  $g \in \text{supp } \nu_k^{(n)}$ , and  $k \geq \hat{k}_r$ . The case  $r = 1$  of (14.5) was done in (10.4), so let us consider  $r > 1$ . By induction,

$$(14.6) \quad -\log \psi_{r-1}(gx_-, gx'_-; \vec{\omega}_-) \leq -\log \psi_{r-1}(x_-, x'_-; \vec{\omega}_-) + C''_{r-1} n.$$

Replacing this and (14.4) in the definition (13.3), we find that

$$\begin{aligned}
-\log \psi_r(x, x'; \vec{\omega}) &= -\beta_{r-1} \log \psi_{r-1}(gx_-, gx'_-; \vec{\omega}_-) \\
&\quad - \log \text{SVP}_r(gx, gx'; \omega_r) \\
&\leq -\beta_{r-1} \log \psi_{r-1}(x_-, x'_-; \vec{\omega}_-) + \beta_{r-1} C''_{r-1} n \\
&\quad - \log \text{SVP}_r(x, x'; \omega_r) + C'_r n \\
&= \psi_r(x, x'; \vec{\omega}) + (\beta_{r-1} C''_{r-1} + C'_r) n.
\end{aligned}$$

Since  $C''_r = C''_{r-1} + C'_r$  and  $\beta_{r-1} < 1$ , this proves (14.5). The estimate remains valid if we exchange the roles of  $x$  and  $x'$ , obviously. Thus, we have also shown that

$$(14.7) \quad -\log \hat{\psi}_r(gx, gx'; \omega) \leq -\log \hat{\psi}_r(x, x'; \vec{\omega}) + C''_r n$$

for any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position,  $g \in \text{supp } \nu_k^{(n)}$ , and  $k \geq \hat{k}_r$ .

Combining (14.7) with (9.1) in the definition (13.24), we get that

$$\begin{aligned}
\Psi_r(gx, gx'; \vec{\omega}) &\leq \log \left( \Omega_r + \hat{\psi}_r(gx, gx'; \vec{\omega})^{-1} \right) \leq \log \left( \Omega_r + e^{C''_r n} \hat{\psi}_r(x, x'; \vec{\omega})^{-1} \right) \\
&\leq \log \left( \Omega_r + \hat{\psi}_r(x, x'; \vec{\omega})^{-1} \right) + C''_r n = \Psi_r(x, x'; \vec{\omega}) + C''_r n,
\end{aligned}$$

as stated.  $\square$

Next, we prove the following extension of Lemma 12.2:

**Lemma 14.3.** *Given  $\delta > 0$ ,  $n \geq N_r$  and  $\omega_r > 0$ , there is  $k'_r = k'_r(\nu_{\mathcal{O}}, \delta, n, \omega_r) \in \mathbb{N}$  and for any  $x, x' \in \mathcal{F}(r, d)$  with  $F'_1 \nsubseteq F_r$  and  $\text{VA}_r(x, x') \geq \omega_r$  there is  $\mathcal{E}''_{k, r} = \mathcal{E}''_{k, r}(\nu_{\mathcal{O}}, \delta, n, x, x', \omega_r) \subset \text{supp } \nu_k^{(n)}$  with  $\nu_k^{(n)}((\mathcal{E}''_{k, r})^c) < \delta$  and*

$$(14.8) \quad -\log \text{SVA}_r(gx, gx'; \omega_r) \leq \max\{-\log \text{SVA}_r(x, x'; \omega_r) - \tilde{\kappa}_r n, \tilde{\theta}_r\}$$

for every  $g \in \mathcal{E}''_{k, r}$  and  $k \geq k'_r$ .

*Proof.* Fix  $\delta > 0$  and  $n \geq N_r$  and  $\omega_r > 0$ . Let  $P_{\omega_r}$  denote the (compact) subset of all  $v \in P$  such that  $\|v^\perp\|/\|v\| \geq \omega_r/2$ . For  $v \in P_{\omega_r}$  and  $g$  in some compact neighborhood  $V_{\omega_r}$  of  $\text{supp } \nu_{\mathcal{O}}^{(n)}$ , consider

$$(14.9) \quad (v, g) \mapsto -\log \frac{\|(gv)^\perp\|}{\|gv\|}.$$

As long as  $V_{\omega_r}$  is sufficiently small, depending on  $\nu_{\mathcal{O}}$ ,  $n$  and  $\omega_r$ , the map (14.9) is well defined and (uniformly) continuous. So, there exists  $\alpha = \alpha(\nu_{\mathcal{O}}, n, \omega_r) > 0$  such that

$$(14.10) \quad -\log \frac{\|(gu)^\perp\|}{\|gu\|} \leq -\log \frac{\|(fv)^\perp\|}{\|fv\|} + \log 2$$

whenever  $d(u, v) < \alpha$  and  $d(g, f) < \alpha$ . Reducing  $\alpha$  if necessary, depending only on  $\omega_r$ , we may also assume that

$$(14.11) \quad d(u, v) < \alpha \quad \Rightarrow \quad -\log \frac{\|u^\perp\|}{\|u\|} \geq -\log \frac{\|v^\perp\|}{\|v\|} - \log 2.$$

Fix  $v_1, \dots, v_l \in P_{\omega_r}$  such that  $P_{\omega_r} \subset B(v_1, \alpha) \cup \dots \cup B(v_l, \alpha)$ . For each  $v \in P_{\omega_r}$  choose  $j \in \{1, \dots, l\}$  such that  $v \in B(v_j, \alpha)$  and define  $\mathcal{E}_{k, 0} = \mathcal{E}_{k, 0}(\nu_{\mathcal{O}}, \delta, n, v, \omega_r) \subset$

supp  $\nu_k^{(n)}$  by

$$(14.12) \quad \mathcal{E}_{k,0} = [\alpha\text{-neighborhood of } \mathcal{E}_0(\nu_\infty, \delta, n, v_j^\perp)] \cap \text{supp } \nu_k^{(n)},$$

where  $\mathcal{E}_0(\nu_\infty, \delta, n, v_j^\perp) \subset \text{supp } \nu_\infty^{(n)}$  is as defined in Proposition 5.5. Since  $\nu_k^{(n)}$  converges to  $\nu_\infty^{(n)}$  in the weak\* topology, the limit inferior of the  $\nu_k^{(n)}$ -measure of (14.12) as  $k \rightarrow \infty$  is greater than or equal to

$$\nu_\infty^{(n)}(\mathcal{E}_0(\nu_\infty, \delta, n, v_j^\perp)) > 1 - \delta$$

for every  $j = 1, \dots, l$ . In particular, there is  $k'_r = k'_r(\nu_\infty, \delta, n, \omega_r) \in \mathbb{N}$  such that

$$(14.13) \quad \nu_k^{(n)}(\mathcal{E}_{k,0}) > 1 - \delta \text{ for every } k \geq k'_r \text{ and } v \in P_{\omega_r}.$$

Given  $x, x' \in \mathcal{F}(r, d)$  with  $F_1' \not\subset F_r$  and  $\text{VA}_r(x, x') \geq \omega_r$ , take  $u = u^E + u^\perp$  to be a non-zero vector that realizes the supremum in the definition (12.2). Then,

$$(14.14) \quad \omega_r \leq \text{VA}_r(x, x') = \frac{\|u^\perp\|}{\|u\|} \quad (\text{in particular, } u \in P_{\omega_r}) \text{ and}$$

$$(14.15) \quad \text{VA}_r(gx, gx') \geq \frac{\|(gu)^\perp\|}{\|gu\|} \geq \text{VA}_r(x, x') \frac{\|(gu)^\perp\|}{\|gu\|} \frac{\|u\|}{\|u^\perp\|}$$

for any  $g \in G$ . Then define

$$(14.16) \quad \mathcal{E}_{k,r}'' = \mathcal{E}_{k,0}(\nu_\infty, \delta, n, u, \omega_r).$$

It follows from (14.13) that  $\nu_k^{(n)}((\mathcal{E}_{k,r}'')^c) < \delta$  for every  $k \geq k'_r$ .

Let  $g \in \mathcal{E}_{k,r}''$  and  $k \geq k'_r$ . Then, by definition, there exist  $v = v^E + v^\perp$  in  $P_{\omega_r}$  (take  $v = v_j$  as in (14.12)) and  $f \in \mathcal{E}_0(\nu_\infty, \delta, n, v^\perp) \subset \text{supp } \nu_\infty^{(n)}$  such that  $d(u, v) < \alpha$  and  $d(g, f) < \alpha$ . Thus, substituting (14.10) and (14.11) in (14.15), we find that

$$(14.17) \quad \text{VA}_r(gx, gx') \geq \frac{1}{2} \frac{\|f^\perp v^\perp\|}{\|fv\|} \geq \frac{1}{4} \text{VA}_r(x, x') \frac{\|f^\perp v^\perp\|}{\|fv\|} \frac{\|v\|}{\|v^\perp\|}.$$

Let  $\tau_0 = \tau_0(\nu_\infty, \delta) > 0$  be as in Proposition 5.5. If  $\|f^\perp v^\perp\|/\|fv\| \geq \tau_0/2$  then the first part of (14.17) gives that (recall (12.9) also)

$$(14.18) \quad -\log \text{VA}_r(gx, gx') \leq -\log \frac{\tau_0}{4} \leq \tilde{\theta}_r.$$

Now suppose that  $\|f^\perp v^\perp\|/\|fv\| < \tau_0/2$ . Then part (2) of Proposition 5.5 gives that

$$(14.19) \quad \frac{\|fv^\perp\|}{\|fv\|} < \frac{1}{2} \text{ and so } \frac{\|fv^E\|}{\|fv\|} > \frac{1}{2}.$$

Substituting (14.19) and  $\|v\| \geq \|v^E\|$  in (14.17), we find that

$$(14.20) \quad \begin{aligned} \text{VA}_r(gx, gx') &\geq \frac{1}{8} \text{VA}_r(x, x') \frac{\|f^\perp v^\perp\|}{\|v^\perp\|} \frac{\|v^E\|}{\|fv^E\|} \\ &= \frac{1}{8} \text{VA}_r(x, x') \frac{\|Df_{v^E}^\perp v^\perp\|}{\|v^\perp\|}. \end{aligned}$$

By part (1) of Proposition 5.5 and the choices of  $\tilde{\kappa}_r > 0$  and  $\tilde{N}_r \in \mathbb{N}$  in (12.9), this implies

$$(14.21) \quad \begin{aligned} -\log \text{VA}_r(gx, gx') &\leq -\log \text{VA}_r(x, x') + \log 8 - 2\tilde{\kappa}_r n \\ &\leq -\log \text{VA}_r(x, x') - \tilde{\kappa}_r n. \end{aligned}$$



The conclusion of the lemma is contained in (14.18) and (14.21).  $\square$

Next, let us prove the following extension of Lemma 12.5:

**Lemma 14.4.** *Given  $\delta > 0$  and  $n \geq N_r$  there exists  $\hat{k}_r = \hat{k}_r(\nu_\infty, \delta, n) \in \mathbb{N}$  such that for any  $x, x' \in E_r^\circ(\varepsilon_r)$  with  $F'_1 \nsubseteq F_r$  and  $-\log \text{VA}_r(x, x') \leq \tilde{\theta}_r + \tilde{\kappa}_r n$  there exists  $\mathcal{E}_{k,r}'' = \mathcal{E}_{k,r}''(\nu_\infty, \delta, n, x, x') \subset \text{supp } \nu_k^{(n)}$  with  $\nu_k^{(n)}((\mathcal{E}_{k,r}'')^c) < \delta$  and*

$$(14.22) \quad -\log d(gF_r, g(F_{r-1} + F'_1)) \leq -\log d(F_r, F_{r-1} + F'_1) - \log \text{VA}_r(x, x') - \hat{\kappa}_r n$$

for every  $g \in \mathcal{E}_{k,r}''$  and  $k \geq \hat{k}_r$ .

*Proof.* Let  $\hat{P}$  denote the (compact) subset of pairs  $(v, w) \in P \times P$  such that

$$(14.23) \quad \frac{\|v^\perp\|}{\|v\|} \leq \hat{\rho}_1 < 2\hat{\rho}_1 \leq \frac{\|w^\perp\|}{\|w\|}$$

( $v^\perp$  and  $w^\perp$  denote the components of  $v$  and  $w$  orthogonal to the equator). Let  $\hat{V}$  be some compact neighborhood of the support of  $\nu_\infty^{(n)}$ . Since,  $\hat{\rho}_r = \hat{\rho}_r(\nu_\infty, n)$ , both  $\hat{P}$  and  $\hat{V}$  depend only on  $\nu_\infty$  and  $n$ . Condition (14.23) ensures that the angle between  $v$  and  $w$  is bounded away from zero and, consequently, so is the angle between  $gv$  and  $gw$  for any  $g \in \hat{V}$ ; both bounds depend only on  $\nu_\infty$  and  $n$ . Thus, the map

$$(14.24) \quad (v, w, g) \mapsto -\log \frac{\|\Pi_{gv}gw\|}{\|gv\|} \frac{\|v\|}{\|w\|}$$

is well-defined and (uniformly) continuous on the domain  $(v, w) \in \hat{P}$  and  $g \in \hat{V}$ . In particular, there exists  $\hat{\alpha} = \hat{\alpha}(\nu_\infty, n) > 0$  such that

$$(14.25) \quad -\log \frac{\|\Pi_{gv}gw\|}{\|gv\|} \frac{\|v\|}{\|w\|} \leq -\log \frac{\|\Pi_{fu}fz\|}{\|fu\|} \frac{\|u\|}{\|z\|} + \log 2$$

whenever  $d(v, u) < \hat{\alpha}$  and  $d(z, w) < \hat{\alpha}$  and  $d(g, f) < \hat{\alpha}$ . Reducing  $\hat{\alpha}$  if necessary, depending only on  $\nu_\infty$  and  $n$ , we may also suppose that

$$(14.26) \quad d(z, w) < \hat{\alpha} \quad \Rightarrow \quad -\log \frac{\|z^\perp\|}{\|z\|} \leq -\log \frac{\|w^\perp\|}{\|w\|} + \log 2.$$

Fix points  $(v_1, w_1), \dots, (v_l, w_l) \in \hat{P}$  such that the balls of radius  $\hat{\rho}$  around these points cover  $\hat{P}$ . For each  $(v, w) \in \hat{P}$  choose  $j \in \{1, \dots, l\}$  such that  $v \in B(v_j, \hat{\alpha})$  and  $w \in B(w_j, \hat{\alpha})$  and then define  $\hat{\mathcal{E}}_{k,0} = \hat{\mathcal{E}}_{k,0}(\nu_\infty, \delta, n, v, w) \subset G$  by

$$(14.27) \quad \hat{\mathcal{E}}_{k,0} = [\hat{\alpha}\text{-neighborhood of } \mathcal{E}_0(\nu_\infty, \delta, n, w_j^\perp)] \cap \text{supp } \nu_k^{(n)},$$

where  $\mathcal{E}_0(\nu_\infty, \delta, n, w_j^\perp)$  is given by Proposition 5.5. Since  $\nu_k^{(n)} \rightarrow \nu_\infty^{(n)}$  in the weak\* topology, the limit inferior of the  $\nu_k^{(n)}$ -measure of (14.27) as  $k \rightarrow \infty$  is greater than or equal to

$$\nu_\infty^{(n)}(\mathcal{E}_0(\nu_\infty, \delta, n, w_j^\perp)) > 1 - \delta$$

for every  $j = 1, \dots, l$ . In particular, there is  $\hat{k}_r = \hat{k}_r(\nu_\infty, \delta, n) \in \mathbb{N}$  such that

$$(14.28) \quad \nu_k^{(n)}(\hat{\mathcal{E}}_{k,0}) > 1 - \delta \text{ for every } k \geq \hat{k}_r \text{ and } (v, w) \in \hat{P}.$$

Given  $x, x' \in E_r^\circ(\varepsilon_r)$  with  $F_1' \not\subset F_r$ , let  $v \in F_r$ ,  $v' \in F_{r-1} + F_1'$  and  $w = v - v'$  be as in (12.15) and (12.16). Just as in (12.21),

$$(14.29) \quad -\log d(gF_r, g(F_{r-1} + F_1')) \leq -\log d(F_r, F_{r-1} + F_1') - \log \frac{\|\Pi_{gv}gw\|}{\|gv\|} \frac{\|v\|}{\|w\|}$$

for any  $g \in G$  such that  $\|gv\| \geq \|gv'\|$  (the case  $\|gv\| \leq \|gv'\|$  is analogous, reversing the roles of  $F_r$  and  $F_{r-1} + F_1'$ ). The assumption  $x \in E_r(\varepsilon_r)$  implies that

$$(14.30) \quad \frac{\|v^\perp\|}{\|v\|} \leq d(x, E) \leq \varepsilon_r < \hat{\rho}_r.$$

Now assume that  $-\log \text{VA}_r(x, x') \leq \tilde{\theta}_r + \tilde{\kappa}_r n$ . Then, using (12.19) and (12.22),

$$(14.31) \quad \frac{\|w^\perp\|}{\|w\|} > \frac{1}{4} \text{VA}_r(x, x') \geq \frac{1}{4} e^{-\tilde{\theta}_r - \tilde{\kappa}_r n} > 2\hat{\rho}_r.$$

Thus,  $(v, w) \in \hat{P}$ . Then define

$$(14.32) \quad \mathcal{E}_{k,r}'' = \hat{\mathcal{E}}_{k,0}(\nu_\infty, \delta, n, v, w).$$

It follows from (14.28) that  $\nu_k^{(n)}((\mathcal{E}_{k,r}'')^c) < \delta$  for every  $k \geq \hat{k}_r$ .

Take  $u = v_j$  and  $z = w_j$  as in (14.27). By definition,  $(u, z) \in \hat{P}$  and  $d(u, v) < \hat{\alpha}$  and  $d(z, w) < \hat{\alpha}$ . Let  $g \in \mathcal{E}_{k,r}''$  and  $k \geq \hat{k}_r$ . The definitions (14.27) and (14.32) imply that there exists  $f \in \mathcal{E}_0(\nu_\infty, \delta, n, z^\perp) \subset \text{supp } \nu_\infty^{(n)}$  such that  $d(g, f) < \theta$ . Combining (14.29) with (14.25),

$$-\log d(gF_r, g(F_{r-1} + F_1')) \leq -\log d(F_r, F_{r-1} + F_1') + \log 2 - \log \frac{\|\Pi_{fu}fz\|}{\|fu\|} \frac{\|u\|}{\|z\|}.$$

Replacing  $z, w$  and  $g$  with  $u, z$  and  $f$  in (12.23) and (12.24), we get that

$$\frac{\|fu\|}{\|u\|} \geq \frac{1}{2} \frac{\|fu^E\|}{\|u^E\|} \quad \text{and} \quad \|\Pi_{fu}fz\| \geq \frac{1}{2} \|f^\perp z^\perp\|.$$

Substituting this in the previous inequality,

$$-\log d(gF_r, g(F_{r-1} + F_1')) \leq -\log d(F_r, F_{r-1} + F_1') + \log 8 - \log \frac{\|f^\perp z^\perp\|}{\|z\|} \frac{\|u^E\|}{\|fu^E\|}.$$

Conditions (14.26) and (14.31) give that

$$-\log \frac{\|z^\perp\|}{\|z\|} \leq -\log \frac{\|w^\perp\|}{\|w\|} + \log 2 \leq -\log \text{VA}_r(x, x') + \log 8.$$

Combining this with the previous inequality, we find that

$$\begin{aligned} -\log d(gF_r, g(F_{r-1} + F_1')) &\leq -\log d(F_r, F_{r-1} + F_1') - \log \text{VA}_r(x, x') \\ &\quad + \log 64 - \log \frac{\|f^\perp z^\perp\|}{\|z^\perp\|} \frac{\|u^E\|}{\|fu^E\|} \\ &= -\log d(F_r, F_{r-1} + F_1') - \log \text{VA}_r(x, x') \\ &\quad + \log 64 - \log \frac{\|Df_{u^E}^\perp z^\perp\|}{\|z^\perp\|}. \end{aligned}$$

By part (1) of Proposition 5.5, and the choice of  $\widehat{N}_r \in \mathbb{N}$  in (12.19), this implies that

$$\begin{aligned} -\log d(gF_r, g(F_{r-1} + F'_1)) &\leq -\log d(F_r, F_{r-1} + F'_1) - \log \text{VA}_r(x, x') + 5 - 2\widehat{\kappa}_r n \\ &\leq -\log d(F_r, F_{r-1} + F'_1) - \log \text{VA}_r(x, x') - \widehat{\kappa}_r n, \end{aligned}$$

as claimed.  $\square$

We deduce the following extension of Proposition 12.3:

**Lemma 14.5.** *For  $\delta > 0$ ,  $n \geq N_r$  and  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$  there is  $k_r''' = k_r'''(\nu_{\mathcal{O}}, \delta, n, \omega_r) \in \mathbb{N}$  and for  $x, x' \in E_r^\circ(\varepsilon_r)$  with  $F'_1 \not\subset F_r$  and  $\text{VA}_r(x, x') \geq \omega_r$  there is  $\mathcal{E}_{k,r}''' = \mathcal{E}_{k,r}'''(\nu_{\mathcal{O}}, \delta, n, x, x', \vec{\omega}) \subset \text{supp } \nu_k^{(n)}$  such that  $\nu_k^{(n)}((\mathcal{E}_{k,r}''')^c) < \delta$  and*

$$(14.33) \quad -\log \psi_r(gx, gx'; \vec{\omega}) \leq -\log \psi_r(x, x'; \vec{\omega}) - \kappa_r'' n.$$

for any  $g \in \mathcal{E}_{k,r}'''$  and  $k \geq k_r'''$ .

*Proof.* Fix  $\delta > 0$  and  $n \geq N_r$  and  $\vec{\omega} \in \mathbb{R}_+^d$ . Recall that

$$N_r > \max \{ \widetilde{N}_r(\nu_{\mathcal{O}}, \delta/2), \widehat{N}_r(\nu_{\mathcal{O}}, \delta/2), 2\widetilde{\theta}_r/(\gamma\widehat{\kappa}_r) \},$$

by (12.25). Define

$$(14.34) \quad \begin{aligned} k_r''' &= \max \{ k_r'(\nu_{\mathcal{O}}, \delta/2, n, \omega), \widehat{k}_r(\nu_{\mathcal{O}}, \delta/2, n) \} \\ \text{and } \mathcal{E}_{k,r}''' &= \mathcal{E}_{k,r}''(\nu_{\mathcal{O}}, \delta/2, n, x, x', \omega) \cap \mathcal{E}_{k,r}''(\nu_{\mathcal{O}}, \delta/2, x, x', n). \end{aligned}$$

By construction,  $\mathcal{E}_{k,r}'''$  is contained in the support of  $\nu_k^{(n)}$  and  $\nu_k^{(n)}((\mathcal{E}_{k,r}''')^c) < \delta$ .

We claim that given any  $x, x' \in E_r^\circ(\varepsilon_r)$  with  $F'_1 \not\subset F_r$  and  $\text{VA}_r(x, x') \geq \omega_r$ ,

$$(14.35) \quad -\log \text{SVP}_r(gx, gx'; \omega_r) \leq -\log \text{SVP}_r(x, x'; \omega_r) - \kappa_r' n$$

for every  $g \in \mathcal{E}_{k,r}'''$  and  $k \geq k_r'''$ . As observed in (13.10), the assumption on  $(x, x')$  implies that  $\text{SVA}_r(x, x'; \omega_r) = \text{VA}_r(x, x')$  and  $\text{SVA}_r(x, x'; \omega_r) = \text{VA}_r(gx, gx')$ , and so (14.35) may be rewritten as

$$(14.36) \quad -\log \text{VP}_r(gx, gx') \leq -\log \text{VP}_r(x, x') - \kappa_r' n.$$

Let  $g \in \mathcal{E}_{k,r}'''$  and  $k \geq k_r'''$ . Suppose first that  $-\log \text{VA}_r(x, x') \geq \widetilde{\theta}_r + \widetilde{\kappa}_r n$ . Then, by Lemma 14.3,

$$(14.37) \quad -\log \text{VA}_r(gx, gx') \leq -\log \text{VA}_r(x, x') - \widetilde{\kappa}_r n.$$

Substituting (14.37) and (8.2) in the definition (12.13) we find that

$$(14.38) \quad \begin{aligned} -\log \text{VP}_r(gx, gx') &\leq -\log \text{VA}_r(x, x') - \widetilde{\kappa}_r n \\ &\quad - \gamma_r \log d(F_r, F_{r-1} + F'_1) + \gamma_r A n \\ &\leq -\log \text{VP}_r(x, x') - \frac{\widetilde{\kappa}_r}{2} n \end{aligned}$$

(we chose  $\gamma_r \leq \widetilde{\kappa}_r/(2A)$  in (12.25)). Now suppose that  $-\log \text{VA}_r(x, x') \leq \widetilde{\theta}_r + \widetilde{\kappa}_r n$ . In this case, Lemma 14.3 yields

$$(14.39) \quad -\log \text{VA}_r(gx, gx') \leq \widetilde{\theta}_r$$

whereas Lemma 14.4 yields

$$(14.40) \quad -\log d(gx, gx') \leq -\log d(x, x') - \log \text{VA}_r(x, x') - \widehat{\kappa}_r n.$$

Substituting (14.39) and (14.40) in the definition (12.13), we obtain

$$\begin{aligned} -\log \text{VP}_r(gx, gx') &\leq -\gamma_r \log d(F_r, F_{r-1} + F'_1) - \gamma_r \log \text{VA}_r(x, x') - \gamma_r \hat{\kappa}_r n \\ &\leq -\log \text{VP}_r(x, x') + \tilde{\theta}_r + (1 - \gamma_r) \log \text{VA}_r(x, x') - \gamma_r \hat{\kappa}_r n. \end{aligned}$$

Since  $\text{VA}_r(x, x') \leq 1$ ,  $\gamma_r \leq 1$ , and  $n \geq N_r \geq 2\tilde{\theta}_r/(\gamma_r \hat{\kappa}_r)$ , it follows that

$$\begin{aligned} (14.41) \quad -\log \text{VP}_r(gx, gx') &\leq -\log \text{VP}_r(x, x') + \tilde{\theta}_r - \gamma_r \hat{\kappa}_r n \\ &\leq -\log \text{VP}_r(x, x') - \frac{\gamma_r \hat{\kappa}_r}{2} n. \end{aligned}$$

Because of the way we chose  $\kappa'_r$  in (12.25), the relations (14.38) and (14.41) contain the claim (14.35).

Now we prove the claim (14.33). The case  $r = 1$  was done in (10.31), so let us suppose  $r > 1$ . By (14.6),

$$-\log \psi_{r-1}(gx_-, gx'_-; \omega) \leq -\log \psi_{r-1}(x_-, x'_-; \vec{\omega}) + C''_{r-1} n.$$

Substituting this and (14.35) in the definition (12.34),

$$\begin{aligned} -\log \psi_r(x, x'; \vec{\omega}) &= -\beta_{r-1} \log \psi_{r-1}(gx_-, gx'_-; \omega) - \log \text{SVP}_r(gx, gx'; \omega_r) \\ &\leq -\beta_{r-1} \log \psi_{r-1}(x_-, x'_-; \vec{\omega}) + \beta_{r-1} C''_{r-1} n \\ &\quad - \log \text{SVP}_r(x, x'; \omega_r) - \kappa'_r n \\ &\leq -\log \psi_r(x, x'; \vec{\omega}) - \kappa''_r, \end{aligned}$$

where the last step uses our choices of  $\beta_{r-1}$ ,  $C''_{r-1}$ , and  $\kappa''_r$  in (12.37) and (12.44).  $\square$

Let us go back to proving Proposition 14.1. Define

$$(14.42) \quad k_r = \max\{\tilde{k}_r(\nu_{\mathcal{O}}, \delta, n), \hat{k}_r(\nu_{\mathcal{O}}, n, \omega_r), k'_r(\nu_{\mathcal{O}}, \delta, n \omega_r), \hat{k}_r(\nu_{\mathcal{O}}, \delta, n), k'''_r(\nu_{\mathcal{O}}, \delta/2, n, \omega_r)\}.$$

Then,  $k_r$  depends only on  $\nu_{\mathcal{O}}$ ,  $\delta$ ,  $n$  and  $\omega_r$ . Part (i) of the proposition is an immediate consequence of Lemma 14.2. To prove part (ii), consider

$$\mathcal{E} = \mathcal{E}'''_{k,r}(\nu_{\mathcal{O}}, \delta/2, n, x, x', \vec{\omega}) \cap \mathcal{E}'''_{k,r}(\nu_{\mathcal{O}}, \delta/2, n, x', x, \vec{\omega})$$

where  $\mathcal{E}'''_{k,r}$  is as given by Lemma 14.5. Then  $\nu_k^{(n)}(\mathcal{E}^c) < \delta$  and

$$(14.43) \quad -\log \hat{\psi}_r(gx, gx'; \vec{\omega}) \leq -\log \hat{\psi}_r(x, x'; \vec{\omega}) - \kappa'_r n \leq -\log \hat{\psi}_r(x, x'; \vec{\omega})$$

for every  $g \in \mathcal{E}$  and  $k \geq k'''_r$ . By (9.2), this implies that

$$(14.44) \quad \Psi_r(gx, gx'; \omega) \leq \Psi_r(x, x'; \omega) \text{ for every } g \in \mathcal{E}.$$

Integrating (14.44) over  $\mathcal{E}$  and (14.1) over the complement, we obtain part (ii).

Next, recall that we took  $\kappa'''_r = \kappa''_r/2 = \kappa'_r/2$  and  $c = e^{-2\kappa'''_r n}$  and  $\varepsilon'''_r > 0$  such that (13.30) holds:

$$-\log \hat{\psi}_r(x, x'; \vec{\omega}) = -\log \hat{\psi}(x, x') \geq \Omega_r/\sqrt{c}$$

for any  $x, x' \in E_r^\circ(\varepsilon'''_r)$  in general position. Then, by (9.3) and the first inequality in (14.43),

$$(14.45) \quad \Psi_r(gx, gx'; \vec{\omega}) \leq \Psi_r(x, x'; \vec{\omega}) + \log \sqrt{c} = \Psi_r(x, x'; \vec{\omega}) - \kappa'''_r n$$

for every  $g \in \mathcal{E}$ . Integrating (14.45) over  $\mathcal{E}$  and (14.1) over the complement, we obtain part (iii) of the proposition.  $\square$

It is clear from the statements of Lemmas 14.2, 14.3 and 14.5 that one may take  $\hat{k}_r$ ,  $k'_r$  and  $k''_r$  to increase to  $\infty$  when  $\omega_r$  decreases to zero and  $\nu_{\mathcal{O}}$ ,  $\delta$ ,  $n$  remain fixed. Then the same is true about the map  $\omega \mapsto k_r(\nu_{\mathcal{O}}, \delta, n, \omega_r)$  defined in (14.42). Hence, we may find  $\check{k}_r = \check{k}_r(\nu_{\mathcal{O}}, \delta, n) \in \mathbb{N}$  and  $\omega_{k,r} = \omega_{k,r}(\nu_{\mathcal{O}}, \delta, n) > 0$  such that

- the sequence  $(\omega_{k,r})_k$  decreases to 0, and
- $k \geq k_r(\nu_{\mathcal{O}}, \delta, n, \omega_{k,r})$  for every  $k \geq \check{k}_r$ .

Denote  $\vec{\omega}_{k,r} = (\omega_{k,1}, \dots, \omega_{k,r})$  and then define

$$(14.46) \quad \hat{\psi}_{k,r}(x, x') = \hat{\psi}_r(x, x'; \vec{\omega}_{k,r}) \text{ and } \Psi_{k,r}(x, x') = \Psi_r(x, x'; \vec{\omega}_{k,r}).$$

## 15. SPREADING OUT

For  $r > 1$  the cut-off in Section 13.2 is insufficient to ensure that  $\Psi_r(x, x'; \vec{\omega})$  is bounded near the border region of  $E_r^\circ(\varepsilon_r)$ . To explain why, let us consider  $x = (F_1, \dots, F_r)$  and  $x' = (F'_1, \dots, F'_r)$  such that  $F_r$  is near the border and  $F'_r$  is far from the border of  $E_r(\varepsilon_r)$ . Then  $-\log \text{VP}_r(x, x')$  is bounded, but the problem is that the term  $-\log \psi_{r-1}(x_-, x'_-)$  may be arbitrarily large, which forces  $\Psi_r(x, x'; \vec{\omega})$  to be arbitrarily large as well. For instance,  $d(F_1, F'_1)$  may be very small, in which case  $-\log \psi_1(F_1, F'_1)$  is very large. Thus, Lemma 9.5 as stated does not extend to  $r > 1$ .

To fix this difficulty, we introduce a ‘‘spreading out’’ Markov operator  $\tilde{\mathcal{Q}}_r$  which leaves  $F_r$  and  $F'_r$  fixed, but averages the function out over the pairs of flags whose  $r$ -dimensional components are  $F_r$  and  $F'_r$ . This is done only on certain domains far from the equator: elsewhere we just take  $\tilde{\mathcal{Q}}_r = \text{id}$ . The main properties of this operator are stated in Proposition 15.1: roughly speaking,  $\tilde{\mathcal{Q}}_r \Psi_r$  is never much bigger than  $\Psi_r$  itself, with equality close to the equator, and it is bounded near the border region. That replaces Lemma 9.5 when  $r > 1$ .

The details of the spreading out construction follow. In Section 16.1 we will incorporate  $\tilde{\mathcal{Q}}_r$  into the definition of our main Markov operators,  $\hat{\mathcal{T}}_{k,r}^{\mathcal{Q}}$ , defined on the space of pairs of flags, and  $\tilde{\mathcal{T}}_{k,r}^{\mathcal{Q}}$ , its lift to the blow-up space  $\mathcal{Y}_r(\varepsilon_r)$ .

**15.1. Homogeneous measures on flag varieties.** The orthogonal group  $O(d)$  acts transitively on the flag space  $\mathcal{F}(r, d)$ . We denote by  $\mu$  the corresponding *homogeneous measure* on  $\mathcal{F}(r, d)$ . This may be described as the image of the Haar probability measure of the orthogonal group under

$$O(d) \rightarrow \mathcal{F}(r, d), \quad g \mapsto gz = (gH_1, \dots, gH_r),$$

for any choice of  $z = (H_1, \dots, H_r)$ , and it is invariant under the  $O(d)$ -action.

Analogously, for any  $F_r \in \text{Gr}(r, d)$ , denote by  $\mu_{F_r}$  the homogeneous measure on

$$(15.1) \quad \mathcal{F}(F_r) = \{(G_1, \dots, G_{r-1}, G_r) \in \mathcal{F}(r, d) : G_r = F_r\}$$

corresponding to the natural action of the orthogonal group  $O(F_r)$  on  $\mathcal{F}(F_r)$ . We also consider

$$(15.2) \quad \mathcal{F}^*(F_r) = \left\{ (G_1, \dots, G_{r-1}, F_r) \in \mathcal{F}(F_r) : d(G_1, E) \geq \frac{1}{2}d(F_r, E) \right\}.$$

It is clear that there exists  $a_r > 0$ , depending only on  $r$ , such that  $\mu_{F_r}(\mathcal{F}^*(F_r)) \geq a_r$ . See Figure 9. We denote by  $\mu_{F_r}^*$  the normalized restriction of  $\mu_{F_r}$  to  $\mathcal{F}^*(F_r)$ .

More generally, there exist  $c_r > 0$  and  $b_r > 0$ , depending only on  $r$ , such that

$$(15.3) \quad \mu_{F_r}^*(\{(G_1, \dots, G_{r-1}, F_r) \in \mathcal{F}^*(F_r) : d(G_1, F) < \rho d(F_r, F)\}) \leq c_r \rho^{b_r}$$

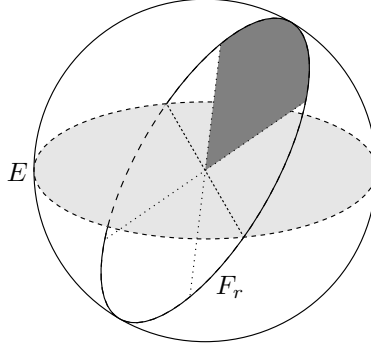


FIGURE 9. The subset  $\mathcal{F}^*(F_r)$  of the flags  $(G_1, \dots, G_{r-1}, F_r)$  such that  $d(G_1, E) \geq d(F_r, E)/2$  is represented (for  $r = 2$  and  $d = 3$ ) by the dark gray region. It corresponds to a definite fraction of all the flags in  $\mathcal{F}(F_r)$ , relative to the homogeneous measure  $\mu_{F_r}$ .

for any  $F \neq F_r$  in  $\text{Gr}(r, d)$  and  $\rho > 0$ . Also (increasing  $c_r$  and decreasing  $b_r$  if necessary),

$$(15.4) \quad \mu_{F_r}^*(\{(G_1, \dots, G_{r-1}, F_r) \in \mathcal{F}^*(F_r) : d(G_1, F_-) < \rho\}) \leq c_r \rho^{b_r}$$

for any  $F_- \in \text{Gr}(r-1, d)$  and  $\rho > 0$  (because  $G_1$  is allowed to vary in a domain whose dimension is strictly greater than  $\dim F_-$ ).

**15.2. Spreading out operators.** Let  $\delta_{(x, x')}$  denote the Dirac mass at a point  $(x, x') \in E_r^\diamond(\varepsilon_r)^2$ , and  $\tau : E_r(\varepsilon_r)^2 \rightarrow [0, 1]$  be a continuous symmetric function such that

$$(15.5) \quad \begin{aligned} \tau &\equiv 1 \text{ on } E_r(\varepsilon_r, \varepsilon'_r) \times E_r(2\varepsilon_r'') \cup E_r(2\varepsilon_r'') \times E_r(\varepsilon_r, \varepsilon'_r) \\ \tau &\equiv 0 \text{ outside } E_r(\varepsilon_r, \varepsilon'_r/2) \times E_r(3\varepsilon_r'') \cup E_r(3\varepsilon_r'') \times E_r(\varepsilon_r, \varepsilon'_r/2). \end{aligned}$$

See Figure 10. Then

$$(15.6) \quad \tilde{q}_{r, x, x'} = (1 - \tau(F_r, F_r'))\delta_{(x, x')} + \tau(F_r, F_r')\mu_{F_r}^* \times \mu_{F_r'}^*,$$

defines a probability measure on  $\mathcal{F}(F_r) \times \mathcal{F}(F_r') \subset E_r^\diamond(\varepsilon_r)^2$  depending continuously on  $(x, x') \in E_r^\diamond(\varepsilon_r)^2$ . The *spreading out operator* is the corresponding continuous Markov operator

$$(15.7) \quad \tilde{Q}_r : \mathcal{B}(E_r^\diamond(\varepsilon_r)^2) \rightarrow \mathcal{B}(E_r^\diamond(\varepsilon_r)^2), \quad \tilde{Q}_r \tilde{\psi}(x, x') = \int_{\mathcal{F}(F_r) \times \mathcal{F}(F_r')} \tilde{\psi} d\tilde{q}_{r, x, x'}.$$

It is clear from the definition that  $\tilde{Q}_r$  is a lift of the identity relative to  $(f, f)$ , meaning that

$$(15.8) \quad \tilde{Q}_r(\tilde{\psi} \circ (f, f)) = \tilde{\psi} \circ (f, f) \text{ for every } \psi \in \mathcal{B}(E_r(\varepsilon_r)^2),$$

where  $f : \mathcal{F}(r, d) \rightarrow \text{Gr}(r, d)$  denotes the *forgetfulness map*

$$(15.9) \quad (F_1, \dots, F_r) \mapsto F_r.$$

Since  $\tau$  is assumed to be symmetric, we also have that  $\tilde{Q}_r$  preserves the space of symmetric functions.

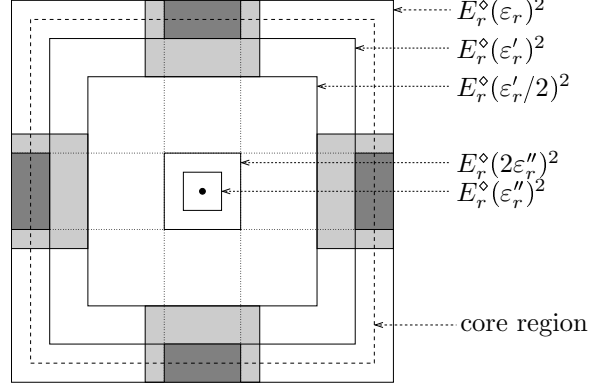


FIGURE 10. Illustrating the spreading out construction. The black dot at the center marks the point  $(E, E)$ . The dashed lined represents the boundary between the  $\nu_k^{(n)}$ -core and the  $\nu_k^{(n)}$ -border of  $E_1(\varepsilon_1)$ . On the dark gray area we do full averaging. No averaging is needed on the white area. On the light gray area we interpolate between the two. In the cut-off region  $E^\diamond(\varepsilon_r, 2\varepsilon_r'')^2$ , where  $\Psi_{k,r} \equiv \log \Omega_r$ , averaging is innocuous.

For each fixed  $\vec{\omega}$ , denote by  $\tilde{Q}_r \Psi_r(\cdot, \cdot; \vec{\omega})$  the image of  $\Psi_r(\cdot, \cdot; \vec{\omega})$  under the operator  $\tilde{Q}_r$ . Since  $\Psi_r(\cdot, \cdot; \vec{\omega})$  is symmetric, by (13.24) and (13.25), the function  $\tilde{Q}_r \Psi_r(\cdot, \cdot; \vec{\omega})$  is also symmetric.

**Proposition 15.1.** *There exist  $K_r = K_r(\nu_\infty) > 0$  and  $\Omega_r = \Omega_r(\nu_\infty, \delta, n) > 1$  such that for every  $(x, x') \in E_r^\diamond(\varepsilon_r)^2$  and  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$  with  $\omega_r < \varepsilon_r'''$ ,*

- (i)  $\tilde{Q}_r \Psi_r(x, x'; \vec{\omega}) = \Psi_r(x, x'; \vec{\omega})$  if  $(x, x') \in E_r^\diamond(\varepsilon_r'/2)^2$ ;
- (ii)  $\tilde{Q}_r \Psi_r(x, x'; \vec{\omega}) \leq \log \Omega_r + K_r \leq \Psi_r(x, x'; \vec{\omega}) + K_r$  if  $(x, x') \notin E_r^\diamond(\varepsilon_r')^2$ .
- (iii)  $\tilde{Q}_r \Psi_r(x, x'; \vec{\omega}) \leq \Psi_r(x, x'; \vec{\omega}) + K_r$  if  $(x, x') \notin E_r^\diamond(\varepsilon_r'/2)^2$ .

The constants  $K_r$  and  $\Omega_r$  are determined in (15.27) below. The proof of this proposition occupies the remainder of the present section. The first step is the following elementary lemma:

**Lemma 15.2.** *Let  $a, b > 0$  and  $(Z, \theta)$  be a probability space. Let  $f : Z \rightarrow (0, \infty)$  be a measurable function such that  $\theta(Z \setminus Z_\tau) \leq \tau$  for every  $\tau \in (0, 1]$ , where  $Z_\tau$  denotes the subset of points  $z \in Z$  such that*

$$(15.10) \quad f(z) \geq a\tau^b.$$

Then

$$\int_Z \log(\Omega + f(z)^{-1}) d\theta(z) \leq \log(\Omega + a^{-1}) + 10b$$

*Proof.* Define  $q(z) = f(z)a^{-1}$ . The assumption means that  $q(z) \geq \tau^b$  for every  $z \in Z_\tau$ . Then, by (9.1) and (9.2),

$$\begin{aligned} \log(\Omega + f(z)^{-1}) &\leq \log(\Omega + a^{-1}) + \max\{0, \log q(z)^{-1}\} \\ &\leq \log(\Omega + a^{-1}) - b \log \tau \end{aligned}$$

for every  $z \in Z_\tau$ . In particular,

$$\int_{Z_\tau \setminus Z_{e\tau}} (\log(\Omega + f(z)^{-1}) - \log(\Omega + a^{-1})) d\theta(x) \leq -e\tau b \log \tau.$$

Considering  $\tau = e^{-j}$  and summing over all  $j$ , we get

$$\int_Z (\log(\Omega + f(z)^{-1}) - \log(\Omega + a^{-1})) d\theta(x) \leq eb \sum_{j=1}^{\infty} j e^{-j},$$

which implies the claim.  $\square$

Keep in mind that we write  $x = (F_1, \dots, F_{r-1}, F_r)$  and  $x' = (F'_1, \dots, F'_{r-1}, F'_r)$ . We use  $y = (G_1, \dots, G_{r-1}, F_r)$  and  $y' = (G'_1, \dots, G'_{r-1}, F'_r)$  to denote the generic elements of  $\mathcal{F}(F_r)$  and  $\mathcal{F}(F'_r)$ , respectively.

**Lemma 15.3.** *There exist  $c'_r = c'_r(\mu_{\mathcal{X}}, \delta, n) > 0$ ,  $c''_r = c''_r(\mu_{\mathcal{X}}, \delta, n) > 0$ ,  $\alpha'_r = \alpha'_r(\nu_{\mathcal{X}}) > 0$ , and  $\alpha''_r = \alpha''_r(\nu_{\mathcal{X}}) > 0$  such that for any  $\tau \in (0, 1]$ ,  $x \in \mathcal{F}(r, d)$ ,  $x' \in E_r^\circ(\varepsilon_r, \varepsilon'_r/2)$ , and  $\vec{\omega} = (\omega_1, \dots, \omega_r)$  with  $\omega_j \in (0, \varepsilon_j''')$  for  $j = 1, \dots, r$ :*

(i) *For every  $y'$  in a set  $X'_\tau \subset \mathcal{F}^*(F'_r)$  with  $\mu_{F'_r}^*(\mathcal{F}^*(F'_r) \setminus X'_\tau) < \tau$ ,*

$$(15.11) \quad -\log \psi_r(x, y'; \vec{\omega}) \leq -\gamma_r \log d(G'_1, F_r) + c'_r - \alpha'_r \log \tau.$$

(ii) *For every  $y'$  in a set  $X''_\tau \subset \mathcal{F}^*(F'_r)$  with  $\mu_{F'_r}^*(\mathcal{F}^*(F'_r) \setminus X''_\tau) < \tau$ ,*

$$(15.12) \quad -\log \psi_r(x, y'; \vec{\omega}) \leq -\gamma_r \log d(F'_r, F_r) + c''_r - \alpha''_r \log \tau.$$

*Proof.* By assumption,  $x' \notin E_r^\circ(\varepsilon'_r/2)$ , that is,  $F'_r \notin E_r(\varepsilon'_r/2)$ . Thus,

$$d(G'_1, E) \geq \frac{1}{2}d(F'_r, E) > \frac{1}{4}\varepsilon'_r$$

for any  $y' \in \mathcal{F}^*(F'_r)$ . Consequently, recalling (12.2),

$$\text{SVA}_r(x, y'; \omega_r) \geq \text{VA}_r(x, y') = d(F_r + G'_1, E) \geq d(G'_1, E) > \frac{1}{4}\varepsilon'_r$$

for any  $y' \in \mathcal{F}(F'_r)$ . Therefore, using (12.13),

$$(15.13) \quad \begin{aligned} -\log \text{SVP}_r(x, y'; \omega_r) &\leq -\log \text{VP}_r(x, y') \\ &\leq -\gamma_r \log d(F_{r-1} + G'_1, F_r) - \log(\varepsilon'_r/4) \\ &\leq -\gamma_r \log d(G'_1, F_r) - \log(\varepsilon'_r/4) \end{aligned}$$

for any  $y' \in \mathcal{F}(F'_r)$  (keep in mind that  $d(\cdot, \cdot)$  is a distance restricted to  $\text{Gr}(r, d)$ ). In particular, (15.13) contains the case  $r = 1$  of the claim (i):

$$(15.14) \quad -\log \psi_1(x, x'; \omega_1) = -\log \text{SVP}_1(x, x'; \omega_1) \leq -\gamma_1 \log d(y', x) + c'_1$$

with  $c'_1 = -\log(\varepsilon'_1/4)$  and  $\alpha'_1 = 0$  (and  $X'_\tau = \mathcal{F}^*(x') = \{x'\}$ ).

Now let  $r > 1$ . By induction, given  $y'_- = (G'_1, \dots, G'_{r-1})$  in  $E_{r-1}^\circ(\varepsilon_{r-1}, \varepsilon'_{r-1}/2)$ , one has

$$(15.15) \quad -\log \psi_{r-1}(x_-, z'; \vec{\omega}_-) \leq -\gamma_{r-1} \log d(H'_1, F_{r-1}) + c'_{r-1} - \alpha'_{r-1} \log(\tau/2)$$

for every  $z' = (H'_1, \dots, H'_{r-2}, G'_{r-1})$  in a set  $Z' \subset \mathcal{F}^*(G'_{r-1})$  with

$$\mu_{G'_{r-1}}^*(\mathcal{F}^*(G'_{r-1}) \setminus Z') < \tau/2.$$

Integrating this estimate over all admissible  $G'_{r-1}$ , we find that

$$(15.16) \quad -\log \psi_{r-1}(x_-, y'_-; \vec{\omega}_-) \leq -\gamma_{r-1} \log d(G'_1, F_{r-1}) + c'_{r-1} - \alpha'_{r-1} \log(\tau/2)$$



for every  $y'$  in a set  $Y' \subset \mathcal{F}^*(F'_r)$  with

$$\mu_{F'_r}^*(\mathcal{F}^*(F'_r) \setminus Y') < \tau/2.$$

On the other hand, taking  $\tau = 2c_r\rho^{b_r}$  and  $F_- = F_{r-1}$  in (15.4), we find that

$$(15.17) \quad -\log d(G'_1, F_{r-1}) \leq \frac{1}{b_r} (-\log \tau + \log c_r)$$

for every  $y'$  in a set  $Y'' \subset \mathcal{F}^*(F'_r)$  with

$$\mu_{F'_r}^*(\mathcal{F}^*(F'_r) \setminus Y'') < \tau/2.$$

Define  $X'_r = Y' \cap Y''$ . Then  $\mu_{F'_r}^*(\mathcal{F}^*(F'_r) \setminus X'_r) < \tau$  and, by (15.16) and (15.17),

$$(15.18) \quad -\log \psi_{r-1}(x_-, y'_-; \vec{\omega}_-) \leq \tilde{c}_r - \tilde{\alpha}_r \log \tau$$

for every  $y' \in X'_r$ , with

$$(15.19) \quad \tilde{c}_r = \frac{\gamma_{r-1}}{b_r} \log c_r + c'_{r-1} + \alpha'_r \log 2 \text{ and } \tilde{\alpha}_r = \frac{\gamma_{r-1}}{b_r} + \alpha'_{r-1}.$$

Combining (15.13) and (15.18), we find that

$$(15.20) \quad \begin{aligned} -\log \psi_r(x, y'; \vec{\omega}) &= -\beta_{r-1} \log \psi_{r-1}(x_-, y'_-; \vec{\omega}_-) - \log \text{SVP}_r(x, y'; \omega_r) \\ &\leq \beta_{r-1} (\tilde{c}_r - \tilde{\alpha}_r \log \tau) - \gamma_r \log d(G'_1, F_r) - \log(\varepsilon'_r/4) \end{aligned}$$

for every  $y'$  in  $X''$ . This proves the claim (ii), with

$$(15.21) \quad c'_r = \beta_{r-1} \tilde{c}_r - \log(\varepsilon'_r/4) \text{ and } \alpha'_r = \beta_{r-1} \tilde{\alpha}_r.$$

Now we deduce the claim (ii). By part (i), there exists  $X' \subset \mathcal{F}(F'_r)$  such that  $\mu_{F'_r}^*(\mathcal{F}^*(F'_r) \setminus X') < \tau/2$  and

$$(15.22) \quad -\log \psi_r(x, y'; \vec{\omega}) \leq -\gamma_r \log d(G'_1, F_r) + c'_r - \alpha'_r \log(\tau/2).$$

for every  $y' \in X'$ . Taking  $\tau = 2c_r\rho^{b_r}$  and  $F = F'_r$  in (15.3), we get that

$$(15.23) \quad \begin{aligned} -\log d(G'_1, F_r) &\leq -\log d(F'_r, F_r) - \log \rho \\ &= -\log d(F'_r, F_r) + \frac{1}{b_r} (\log 2c_r - \log \tau) \end{aligned}$$

for every  $y'$  in a set  $Y \subset \mathcal{F}(F'_r)$  with  $\mu_r^*(\mathcal{F}^*(F'_r) \setminus Y) < \tau/2$ . Define  $X''_r = X' \cap Y$ . Then  $\mu_{F'_r}^*(\mathcal{F}^*(F'_r) \setminus X''_r) < \tau$  and the relations (15.22) and (15.23) imply that (15.12) holds for every  $y' \in X''_r$ , with

$$(15.24) \quad c''_r = c'_r + \frac{\gamma_r}{b_r} \log 2c_r + \alpha'_r \log 2 \text{ and } \alpha''_r = \alpha_r + \frac{\gamma_r}{b_r}.$$

This completes the proof of the lemma.  $\square$

**Corollary 15.4.** *There exist  $K_r = K_r(\nu_{\mathcal{O}}) > 0$  and  $\Omega_r = \Omega_r(\nu_{\mathcal{O}}, \delta, n) > 1$  such that for all  $(x, x') \in E_r^{\circ}(\varepsilon_r)^2$  such that  $x \notin E_r^{\circ}(\varepsilon'_r/2)$  or  $x' \notin E_r^{\circ}(\varepsilon'_r/2)$ , and for all  $\vec{\omega} = (\omega_1, \dots, \omega_r) \in \mathbb{R}_+^d$  with  $\omega_r < \varepsilon_r''$ ,*

$$(15.25) \quad \int_{\mathcal{F}(F'_r)} \Psi_r(x, y'; \vec{\omega}) d\mu_{F'_r}^*(y') \text{ and } \int_{\mathcal{F}(F_r)} \Psi_r(y, x'; \vec{\omega}) d\mu_{F_r}^*(y)$$

are both bounded above by  $\log \Omega_r + K_r$ .

*Proof.* Since the function  $\Psi_r(\cdot, \cdot; \vec{\omega})$  is symmetric, it suffices to consider the first integral in (15.25). Initially, suppose that  $x \notin E_r^\circ(2\varepsilon_r'')$  and  $x' \notin E_r^\circ(2\varepsilon_r'')$ . Then  $\Psi_r(x, x'; \vec{\omega}) = \log \Omega_r$  and, in fact,  $\Psi_r(\cdot, \cdot; \vec{\omega}) \equiv \log \Omega_r$  on  $\mathcal{F}(F_r) \times \mathcal{F}(F_r')$ . Hence the integral is equal to  $\log \Omega_r$ , and so the claim is trivial in this case.

From now on, let us assume that some of the points  $x$  and  $x'$  is in  $E_r^\circ(2\varepsilon_r'')$ . Since the other is necessarily outside  $E_r^\circ(\varepsilon'/2)$ , by hypothesis, it follows that

$$(15.26) \quad d(F_r, F_r') \geq \varepsilon'_r/2 - 2\varepsilon_r''.$$

We are going to apply Lemma 15.2 to

$$\begin{aligned} Z &= \mathcal{F}(F_r'), \quad \theta = \mu_{F_r'}^*, \quad \Omega = \Omega_r, \\ f(y') &= \psi_r(x, y'; \vec{\omega}), \quad a = d(F_r, F_r')^{\gamma_r} e^{-c_r''}, \quad \text{and } b = \alpha_r''. \end{aligned}$$

The assumption (15.10) of Lemma 15.2 corresponds precisely to part (ii) of the conclusion of Lemma 15.3: given any  $\tau \in (0, 1]$ ,

$$-\log \psi_r(x, y'; \vec{\omega}) \leq -\gamma_r \log d(F_r', F_r) + c_r'' - \alpha_r'' \log \tau$$

for every  $y'$  in a set  $Z_r'' \subset Z$  with  $q_{r, x'}(Z \setminus Z_r'') < \tau$ . The conclusion of Lemma 15.2 asserts that

$$\begin{aligned} \int_{\mathcal{F}(F_r')} \Psi_r(x, y'; \vec{\omega}) d\mu_{F_r'}^*(y') &= \int_{\mathcal{F}(F_r')} \log(\Omega_r + \psi_r(x, y'; \vec{\omega})^{-1}) d\mu_{F_r'}^*(y') \\ &\leq \log(\Omega_r + e^{c_r''} d(F_r, F_r')^{-\gamma_r}) + 10\alpha_r'' \end{aligned}$$

Define

$$(15.27) \quad \Omega_r = e^{c_r''} (\varepsilon'_r/2 - 2\varepsilon_r'')^{-\gamma_r} \quad \text{and } K_r = 10\alpha_r'' + \log 2.$$

Replacing (15.26) in the previous inequality we find that

$$\begin{aligned} \int_{\mathcal{F}(F_r')} \Psi_r(x, y'; \vec{\omega}) d\mu_{F_r'}^*(y') &\leq \log(\Omega_r + e^{c_r''} (\varepsilon'_r/2 - 2\varepsilon_r'')^{-\gamma_r}) + 10\alpha_r'' \\ &\leq \log \Omega_r + K_r \end{aligned}$$

as claimed.  $\square$

*Proof of Proposition 15.1.* It follows immediately from the definitions (15.5) and (15.6) that  $\tilde{\mathcal{Q}}_r \Psi_r = \Psi_r$  outside  $E_r^\circ(\varepsilon_r, \varepsilon'_r/2) \times E_r^\circ(3\varepsilon_r'') \cup E_r^\circ(3\varepsilon_r'') \times E_r^\circ(\varepsilon_r, \varepsilon'_r/2)$ . This contains part (i) of the proposition. We also have that  $\tilde{\mathcal{Q}}_r \Psi_r = \Psi_r = \log \Omega_r$  on the cut-off region  $E_r^\circ(\varepsilon_r, 2\varepsilon_r'')^2$ . Thus (check Figure 10), to complete the proof of part (ii) we only have to consider the case when  $(x, x') \in E_r^\circ(\varepsilon_r, \varepsilon'_r) \times E_r^\circ(2\varepsilon_r'') \cup E_r^\circ(2\varepsilon_r'') \times E_r^\circ(\varepsilon_r, \varepsilon'_r)$ . In this case  $\tilde{q}_{r, r, x'} = \mu_{F_r}^* \times \mu_{F_r'}^*$ , and so Corollary 15.4 gives that

$$\begin{aligned} \tilde{\mathcal{Q}}_r \Psi_r(x, x'; \omega) &= \int_{\mathcal{F}(F_r)} \int_{\mathcal{F}(F_r')} \Psi_r(u, u'; \vec{\omega}) d\mu_{F_r'}^*(u') d\mu_{F_r}^*(u) \\ &\leq \int_{\mathcal{F}(F_r)} (\log \Omega_r + K_r) d\mu_{F_r}^*(u) = \log \Omega_r + K_r. \end{aligned}$$

This proves the first inequality in part (ii). The second one is an immediate consequence of the definition of  $\Psi_r$ .

We are left to proving part (iii) of the proposition. Combining the definition (15.6) with Corollary 15.4, we see that

$$(15.28) \quad \begin{aligned} \int_{\mathcal{F}(F_r) \times \mathcal{F}(F'_r)} \Psi_r(u, u'; \vec{\omega}) d\tilde{q}_{r,x,x'}(u, u') \\ \leq (1 - \tau(F_r, F'_r)) \Psi_r(x, x'; \vec{\omega}) + \tau(F_r, F'_r) (\log \Omega_r + K_r) \\ \leq \Psi_r(x, x'; \vec{\omega}) + \tau(F_r, F'_r) K_r \leq \Psi_r(x, x'; \vec{\omega}) + K_r, \end{aligned}$$

as claimed.  $\square$

**Corollary 15.5.** *For any  $\delta > 0$ ,  $n \geq N_r$ , the following holds for any  $k \geq \check{k}_r$ :*

(i) *For any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position with  $\Psi_{k,r}(x, x') > \log \Omega_r$ ,*

$$\int_G \tilde{Q}_r \Psi_{k,r}(gx, gx') d\nu_k^{(n)}(g) \leq \Psi_{k,r}(x, x') + C_r'' n + K_r.$$

(ii) *For any  $x, x' \in E_r^\circ(\varepsilon_r)$  in general position satisfying  $\Psi_{k,r}(x, x') > \log \Omega_r$ ,  $\text{VA}_r(x, x') \geq \omega_{k,r}$ , and  $\text{VA}_r(x', x) \geq \omega_{k,r}$ ,*

$$\int_G \tilde{Q}_r \Psi_{k,r}(gx, gx') d\nu_k^{(n)}(g) \leq \Psi_{k,r}(x, x') + C_r'' \delta n + K_r.$$

(iii) *For any  $x, x' \in E_r^\circ(\varepsilon_r''')$  in general position satisfying  $\text{VA}_r(x, x') \geq \omega_{k,r}$  and  $\text{VA}_r(x', x) \geq \omega_{k,r}$ ,*

$$\int_G \tilde{Q}_r \Psi_{k,r}(gx, gx') d\nu_k^{(n)}(g) \leq \Psi_{k,r}(x, x') - (\kappa_r'' - C_r'' \delta)n.$$

*Proof.* Parts (i) and (iii) of Proposition 15.1 imply that

$$\tilde{Q}_r \Psi_r(x, x') \leq \Psi_r(x, x') + K_r \text{ for every } (x, x') \in E_r^\circ(\varepsilon_r)^2,$$

Then the claims in parts (i) and (ii) of the corollary follow immediately from the corresponding statements in Proposition 14.1. In the context of part (iii) of the corollary, we even have that  $\tilde{Q}_r \Psi_r(x, x') = \Psi_r(x, x')$ , and so the claim corresponds exactly to part (iii) of Proposition 14.1.  $\square$

## 16. RECOUPLING AND CONCLUSION

By induction, there are constants  $\varepsilon_r > 0$  and  $n_r \in \mathbb{N}$ , continuous Markov operators

$$(16.1) \quad \mathcal{T}_{k,r} : \mathcal{B}(E_r(\varepsilon_r)) \rightarrow \mathcal{B}(E_r(\varepsilon_r)), \quad \mathcal{T}_{k,r} \varphi(F_r) = \int_{E_r(\varepsilon_r)} \psi d\sigma_{k,r,F_r}$$

adapted to  $(\nu_k, E_r(\varepsilon_r))$ , and  $\mathcal{T}_{k,r}$ -invariant probability measures  $\eta_{k,r}$  on  $E_r(\varepsilon_r)$  such that the sequence  $\eta_{\mathcal{O},r} = \lim_k \eta_{k,r}$  exists and satisfies  $\eta_{\mathcal{O},r}(E_r) > 0$ . Up to "localizing" the Markov operators as described in Section 6.5, if necessary, we may assume that  $\varepsilon_r > 0$  is small enough that

$$(16.2) \quad \eta_{\mathcal{O},r}(E_r(\varepsilon_r) \setminus E_r) < \frac{1}{10} \eta_{k,r}(E_r).$$

Then, for every  $k$  sufficiently large,

$$(16.3) \quad \eta_{k,r}(E_r(\varepsilon_r, \varepsilon_r''')) < \frac{2}{10} \eta_{k,r}(E_r(\varepsilon_r)).$$

We are going to show that this leads to a contradiction when  $\dim E = r$ , and to recover all this information for  $r + 1$  when  $\dim E > r$ .

**16.1. Markov operators on flag varieties.** Let  $k$  and  $r$  be fixed. We are going to extend  $\mathcal{T}_{k,r}$  to a suitable Markov operator  $\mathcal{T}_{k,r}^\diamond$  in the space of flags, as follows.

The first step is to find a suitable lift of  $\sigma_{k,r,x}$  to a probability measure on the group  $G = \mathrm{GL}(\mathbb{R}^d)$ . For each  $F_r \in \mathrm{Gr}(r, d)$  define

$$\mathfrak{E}_{F_r} : G \rightarrow \mathrm{Gr}(r, d), \quad g \mapsto gF_r.$$

**Lemma 16.1.** *There exists a continuous family  $\{\mu_{k,r,F_r} : F_r \in E_r(\varepsilon_r)\}$  of probability measures on  $G$  such that  $(\mathfrak{E}_{F_r})_*\mu_{k,r,F_r} = \sigma_{k,r,F_r}$  for every  $F_r \in E_r(\varepsilon_r)$ .*

*Proof.* Write  $\mathbb{R}^d = E \oplus E^\perp$ . Every  $F \in E_r(\varepsilon_r)$  is the graph of a linear map  $u_F : E \rightarrow E^\perp$ . Define

$$h_F : \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad v^E + v^\perp \mapsto v^E + (v^\perp + u_F(v^E)).$$

Then  $F \mapsto h_F$  is a continuous injective map from  $E_r(\varepsilon_r)$  to  $G$  with  $h_F(E) = F$ . For each fixed  $F_r \in E_r(\varepsilon_r)$ , define

$$\mathfrak{L}_{F_r} : E_r(\varepsilon_r) \rightarrow G, \quad F \mapsto h_F \circ h_{F_r}^{-1}.$$

Then  $\mathfrak{L}_{F_r}$  is a continuous injection and a right-inverse of  $\mathfrak{E}_{F_r}$ :

$$(16.4) \quad \mathfrak{E}_{F_r}(\mathfrak{L}_{F_r}(F)) = \mathfrak{L}_{F_r}(F)F_r = (h_F \circ h_{F_r}^{-1})(F_r) = h_F(E) = F$$

for every  $F \in E_r(\varepsilon_r)$ . Define  $\mu_{k,r,F_r} = (\mathfrak{L}_{F_r})_*\sigma_{k,r,F_r}$ . It is clear that this varies continuously with  $F_r$ . The claim in the lemma follows directly from (16.4).  $\square$

Since  $\mathcal{T}_{k,r}$  is adapted to  $(\nu_k^{(n)}, E_r(\varepsilon_r))$ , there exists a neighborhood  $V$  of the  $\nu_k^{(n)}$ -core of  $E_r(\varepsilon_r)$  such that  $\sigma_{k,r,F_r}$  coincides with  $\nu_{k,F_r}^{(n)} = (\mathfrak{E}_{F_r})_*\nu_k^{(n)}$  for every  $F_r \in V$ . Let  $\tau : E_r(\varepsilon_r) \rightarrow [0, 1]$  be a continuous function vanishing on a neighborhood  $U \subset V$  of the  $\nu_k^{(n)}$ -core of  $E_r(\varepsilon_r)$ , and constant equal to 1 outside  $U$ . Define

$$\nu_{k,r,F_r} = (1 - \tau(F_r))\nu_k^{(n)} + \tau(F_r)\mu_{k,r,F_r}.$$

Observe that  $(\mathfrak{E}_{F_r})_*\nu_{k,r,F_r} = \sigma_{k,r,F_r}$  for all  $F_r \in E_r(\varepsilon_r)$ , and so the operator  $\mathcal{T}_{k,r} : \mathcal{B}(E_r(\varepsilon_r)) \rightarrow \mathcal{B}(E_r(\varepsilon_r))$  may be rewritten as

$$\mathcal{T}_{k,r}\psi(F_r) = \int_G \psi(gF_r) d\nu_{k,r,F_r}(g).$$

We extend this to

$$(16.5) \quad \mathcal{T}_{k,r}^\diamond : \mathcal{B}(E_r^\diamond(\varepsilon_r)) \rightarrow \mathcal{B}(E_r^\diamond(\varepsilon_r)), \quad \mathcal{T}_{k,r}^\diamond\psi(x) = \int_G \psi(gx) d\nu_{k,r,F_r}(g),$$

where  $x = (F_1, \dots, F_r)$  and  $gx = (gF_1, \dots, gF_r)$ . It is clear that  $\mathcal{T}_{k,r}^\diamond$  projects to  $\mathcal{T}_{k,r}$  under the forgetfulness map (15.9). Let  $\eta_{k,r}^\diamond$  be a  $\mathcal{T}_{k,r}^\diamond$ -invariant probability measure projecting to  $\eta_{k,r}$ .

**16.2. Recoupling.** We move to construct suitable self-couplings for the Markov operators  $\mathcal{T}_{k,r}^\diamond$ . Begin by writing the definition in (16.5) as

$$(16.6) \quad \mathcal{T}_{k,r}^\diamond\psi(x) = \int_{E_r^\diamond(\varepsilon_r)} \psi d\sigma_{k,r,x}^\diamond \text{ for } \psi \in \mathcal{B}(E_r^\diamond(\varepsilon_r)),$$

where  $\sigma_{k,r,x}^\diamond$  is the push-forward of  $\nu_{k,r,F_r}$  under the map  $G \rightarrow \mathcal{F}(r, d)$ ,  $g \mapsto gx$ . By construction,  $\sigma_{k,r,x}^\diamond$  coincides with the push-forward  $\nu_{k,r,x}^{(n)}$  of  $\nu_k^{(n)}$  whenever  $F_r$  is in the neighborhood  $U$  of  $\nu_k^{(n)}$ -core of  $E_r(\varepsilon_r)$ . We denote by  $\nu_{k,r,x,x}^{(n)}$  the push-forward of  $\nu_k^{(n)}$  under the diagonal embedding  $G \rightarrow \mathcal{F}(r, d)^2$ ,  $g \mapsto (gx, gx')$ .

Consider  $X = X' = E_r^\diamond(\varepsilon_r)$ ,  $Y = Y' = E_r^\diamond(\varepsilon_r)$ , and  $\eta_y = \eta'_y = \sigma_{k,r,y}^\diamond$  for every  $y \in Y$ . Let  $K$  be the (compact) subset  $\text{Diag}_r$  of pairs  $(x, x') \in E_r^\diamond(\varepsilon_r)^2$  which are *not* in general position, that is, such that either  $F'_1 \subset F_r$  or  $F_1 \subset F'_r$ . It is clear that  $K(x')$  and  $K'(x)$  are algebraic subvarieties of  $E_r^\diamond(\varepsilon_r)$ , and so

$$\sigma_{k,r,y}^\diamond(K(x')) = \sigma_{k,r,y}^\diamond(K'(x)) = \{0\}$$

for every  $x, x'$  and  $y$ . This means that (6.10) holds in this setting, and so we may use Proposition 6.9 to find a continuous family

$$\left\{ \tilde{\theta}_{k,r,x,x'} : (x, x') \in E_r^\diamond(\varepsilon_r)^2 \right\}$$

of generic probability measures on  $E_r^\diamond(\varepsilon_r)^2$  such that each  $\theta_{k,r,x,x'}$  is a coupling of  $\sigma_{k,r,x}^\diamond$  and  $\sigma_{k,r,x'}^\diamond$  vanishing on a uniform neighborhood of  $\text{Diag}_r$ .

Let  $\tilde{\omega} : E_r^\diamond(\varepsilon_r)^2 \rightarrow [0, 1]$  be a continuous function such that  $\tilde{\omega}(x, x') = 0$  if  $F_r$  and  $F'_r$  are both in  $E_r^\diamond(\tilde{\varepsilon}_r)$  and  $\tilde{\omega}(x, x') = 1$  if either of them is outside  $E_r^\diamond(2\tilde{\varepsilon}_r)$ . Then

$$(16.7) \quad \tilde{\sigma}_{k,r,x,x'}^\diamond = (1 - \tilde{\omega}(x, x')) \nu_{k,r,x,x'}^{(n)} + \tilde{\omega}(x, x') \tilde{\theta}_{k,r,x,x'}$$

is a coupling of  $\sigma_{k,r,x}^\diamond$  and  $\sigma_{k,r,x'}^\diamond$  depending continuously on  $(x, x')$ , and so

$$\tilde{\mathcal{T}}_{k,r}^\diamond : \mathcal{B}(E_r^\diamond(\varepsilon_r)^2) \rightarrow \mathcal{B}(E_r^\diamond(\varepsilon_r)^2), \quad \tilde{\mathcal{T}}_{k,r}^\diamond \tilde{\psi}(x, x') = \int_{E_r^\diamond(\varepsilon_r)^2} \tilde{\psi} d\tilde{\sigma}_{k,r,x,x'}^\diamond,$$

is a continuous self-coupling of  $\mathcal{T}_{k,r}^\diamond$ . Just as we did for  $r = 1$ , we must modify these operators, by recoupling the measures  $\sigma_{k,r,x}^\diamond$  and  $\sigma_{k,r,x'}^\diamond$  in a suitable way on the region  $E_r^\diamond(2\tilde{\varepsilon}_r, \varepsilon_r'')$ .

For  $x \in E_r^\diamond(2\tilde{\varepsilon}_r, \varepsilon_r'')$ , it follows from (9.16) that the subset of  $g \in \text{supp } \nu_k^{(n)}$  such that  $gx \in E_r^\diamond(2\varepsilon_r')$  is disjoint from the set  $\mathcal{D}_k(F_r)$  given by Corollary 5.7. Hence,

$$(16.8) \quad \nu_{k,r,x}^{(n)}(E_r^\diamond(2\varepsilon_r')) \leq \nu_k^{(n)}(\mathcal{D}_k(F_r)^c) < \delta \text{ for every } x \text{ in } E_r^\diamond(2\tilde{\varepsilon}_r, \varepsilon_r'').$$

Note that  $\sigma_{k,r,x}^\diamond = \nu_{k,r,x}^{(n)}$  if  $F_r$  is in the  $\nu_k^{(n)}$ -core of  $E_r(\varepsilon_r)$ .

Take  $X = X' = E_r^\diamond(\varepsilon_r)$ ,  $Y = Y' = E_r^\diamond(2\tilde{\varepsilon}_r, \varepsilon_r'')$ , and  $\eta_y = \eta'_y = \sigma_{k,r,y}^\diamond$  for every  $y \in Y$ . Moreover, let  $K = E_r^\diamond(\varepsilon_r')^2 \cup \text{Diag}_r$ . On the one hand, (16.8) implies that  $\sigma_{k,r,y}^\diamond(E_r^\diamond(\varepsilon_r'))$  and  $\sigma_{k,r,y}^\diamond(E_r^\diamond(\varepsilon_r'))$  are less than  $\delta < 1/2$  for every  $y \in Y$ . On the other hand,

$$\begin{aligned} \text{Diag}_r(x') &= \{x \in E_r^\diamond(\varepsilon_r) : F'_1 \subset F_r \text{ or } F_1 \subset F'_r\} \text{ and} \\ \text{Diag}'_r(x) &= \{x' \in E_r^\diamond(\varepsilon_r) : F'_1 \subset F_r \text{ or } F_1 \subset F'_r\} \end{aligned}$$

are algebraic subvarieties of  $E_r^\diamond(\varepsilon_r)$ , and so they have zero  $\sigma_{k,r,y}^\diamond$ -measure for every  $y \in Y$ . These two observations show that

$$K(x') = E_r^\diamond(\varepsilon_r') \cup \text{Diag}_r(x') \text{ and } K'(x) = E_r^\diamond(\varepsilon_r') \cup \text{Diag}'_r(x)$$

satisfy (6.10). So we may use Proposition 6.9 to find a continuous family

$$\left\{ \zeta_{k,r,x,x'} : (x, x') \in E_r^\diamond(2\tilde{\varepsilon}_r, \varepsilon_r'')^2 \right\}$$

of generic probability measures on  $E_r^\diamond(\varepsilon_r)^2$  such that every  $\zeta_{k,r,x,x'}$  is a coupling of  $\sigma_{k,r,x}^\diamond$  and  $\sigma_{k,r,x'}^\diamond$  vanishing on a uniform neighborhood of  $\text{Diag}_r$  and such that

$$(16.9) \quad \zeta_{k,r,x,x'}(E_r^\diamond(\varepsilon_r')^2) = 0.$$

Fix a continuous function  $\tau : E_r^\diamond(\varepsilon_r)^2 \rightarrow [0, 1]$  such that  $\tau \equiv 1$  on  $E_r^\diamond(\tilde{\varepsilon}_r, 2\varepsilon_r'')$  and  $\tau \equiv 0$  on the complement of  $E_r^\diamond(2\tilde{\varepsilon}_r, \varepsilon_r'')$ , and then define

$$(16.10) \quad \hat{\sigma}_{k,r,x,x'}^\diamond = (1 - \tau(x, x')) \tilde{\sigma}_{k,r,x,x'}^\diamond + \tau(x, x') \zeta_{k,r,x,x'}$$

for every  $(x, x') \in E_r^\diamond(\varepsilon_r)^2$ . Then  $\hat{\sigma}_{k,r,x,x'}^\diamond$  is a coupling of  $\sigma_{k,r,x}^\diamond$  and  $\sigma_{k,r,x'}^\diamond$  depending continuously on  $(x, x')$ , and so

$$\hat{\mathcal{T}}_{k,r}^\diamond : \mathcal{B}(E_r^\diamond(\varepsilon_r)^2) \rightarrow \mathcal{B}(E_r^\diamond(\varepsilon_r)^2), \quad \hat{\mathcal{T}}_{k,r}^\diamond \tilde{\psi}(x, x') = \int_{E_r^\diamond(\varepsilon_r)^2} \tilde{\psi} d\hat{\sigma}_{k,1,x,x'}^\diamond$$

is another continuous self-coupling of  $\mathcal{T}_{k,r}^\diamond$ , coinciding with  $\tilde{\mathcal{T}}_{k,r}^\diamond$  outside the *recoupling region*  $E_r^\diamond(2\tilde{\varepsilon}_r, \varepsilon_r'')$ .

Finally, define  $\hat{\mathcal{T}}_{k,r}^\diamond : \mathcal{B}(E_r^\diamond(\varepsilon_r)^2) \rightarrow \mathcal{B}(E_r^\diamond(\varepsilon_r)^2)$  by  $\hat{\mathcal{T}}_{k,r}^\diamond = \hat{\mathcal{T}}_{k,r}^\diamond \circ \tilde{\mathcal{Q}}_r$ , that is,

$$(16.11) \quad \hat{\mathcal{T}}_{k,r}^\diamond \tilde{\varphi}(x, x') = \int_{E_r^\diamond(\varepsilon_r)^2} (\tilde{\mathcal{Q}}_r \tilde{\varphi}) d\hat{\sigma}_{k,1,x,x'}^\diamond.$$

Let  $\pi_i : E_r(\varepsilon_r)^2 \rightarrow E_r(\varepsilon_r)$  and  $\pi_i^\diamond : E_r^\diamond(\varepsilon_r)^2 \rightarrow E_r^\diamond(\varepsilon_r)$  denote the projections to the  $i$ th factor,  $i = 1, 2$ , and  $f$  be the forgetfulness map (15.9). By (16.11) and (15.8),

$$\hat{\mathcal{T}}_{k,r}^\diamond(\tilde{\psi} \circ (f, f)) = \hat{\mathcal{T}}_{k,r}^\diamond(\tilde{\mathcal{Q}}_r(\tilde{\psi} \circ (f, f))) = \hat{\mathcal{T}}_{k,r}^\diamond(\tilde{\psi} \circ (f, f))$$

for any  $\tilde{\psi} \in \mathcal{B}(E_r(\varepsilon_r)^2)$ . Take  $\tilde{\psi} = \psi \circ \pi_i$  for any  $\psi \in \mathcal{B}(E_r(\varepsilon_r))$ . Observing that  $\pi_i \circ (f, f) = f \circ \pi_i^\diamond$ , and keeping in mind that  $\hat{\mathcal{T}}_{k,r}^\diamond$  is a self-coupling of  $\mathcal{T}_{k,r}^\diamond$  and the latter projects to  $\mathcal{T}_{k,r}$  under the forgetfulness map, we get that

$$(16.12) \quad \hat{\mathcal{T}}_{k,r}^\diamond(\psi \circ f \circ \pi_i^\diamond) = \hat{\mathcal{T}}_{k,r}^\diamond(\psi \circ f \circ \pi_i^\diamond) = \mathcal{T}_{k,r}^\diamond(\psi \circ f) \circ \pi_i^\diamond = (\mathcal{T}_{k,r} \psi) \circ f \circ \pi_i^\diamond.$$

In other words,  $\hat{\mathcal{T}}_{k,r}^\diamond$  projects to  $\mathcal{T}_{k,r}$  under  $f \circ \pi_i^\diamond$  for any  $i = 1, 2$ .

**Remark 16.2.** Unlike  $\hat{\mathcal{T}}_{k,r}^\diamond$ , this  $\hat{\mathcal{T}}_{k,r}^\diamond$  is not a coupling of operators on  $\mathcal{B}(E_r^\diamond(\varepsilon_r))$  because  $\tilde{\mathcal{Q}}_r$  itself is not a coupling. That could be remedied by choosing  $\tilde{\mathcal{Q}}_r$  differently. However, our choice is convenient for lifting these operators to the blow-up space  $\mathcal{Y}_r(\varepsilon_r)$ , as we will see in Section 16.4.

We will need the following extension of Lemma 11.1 to  $r > 1$ :

**Lemma 16.3.** *Let  $(x, x') \in E_r^\diamond(\varepsilon_r)^2$  be such that*

- (a) *either at least one of the points  $x$  or  $x'$  is in the  $\nu_k^{(n)}$ -border of  $E_r^\diamond(\varepsilon_r)$ ,*
- (b) *or both  $x$  and  $x'$  are in the  $\nu_k^{(n)}$ -core of  $E_r^\diamond(\varepsilon_r)$  but outside  $E_r^\diamond(2\varepsilon_r'')$ .*

*Then  $\hat{\sigma}_{k,r,x,x'}^\diamond(E_r^\diamond(\varepsilon_r')^2) = 0$  and  $\hat{\mathcal{T}}_{k,r}^\diamond \Psi_{k,r}(x, x') \leq \log \Omega_r + K_r$ .*

*Proof.* Let us begin by proving the claim that  $\hat{\sigma}_{k,r,x,x'}^\diamond$  vanishes on  $E_r^\diamond(\varepsilon_r')^2$ . If  $x$  is in the  $\nu_k^{(n)}$ -border of  $E_r^\diamond(\varepsilon_r)$  then, using (13.22),

$$\hat{\sigma}_{k,r,x,x'}^\diamond(E_r^\diamond(\varepsilon_r')^2) \leq \sigma_{k,r,x}^\diamond(E_r^\diamond(\varepsilon_r')) \leq \sigma_{k,r,x}^\diamond \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_r^\diamond(\varepsilon_r) \right) = 0.$$

The same argument applies when  $x'$  is in the  $\nu_k^{(n)}$ -border of  $E_r^\diamond(\varepsilon_r)$ . This settles the claim in case (a). Now let  $x$  and  $x'$  be as in (b). Keep in mind that  $\sigma_{k,r,x}^\diamond = \nu_{k,r,x}^{(n)}$  and  $\sigma_{k,r,x'}^\diamond = \nu_{k,r,x'}^{(n)}$ . By (16.9),  $\zeta_{k,r,x,x'}$  vanishes on  $E_r^\diamond(\varepsilon_r')^2$ , and so (16.10) gives that

$$\hat{\sigma}_{k,r,x,x'}^\diamond(E_r^\diamond(\varepsilon_r')^2) = (1 - \tau(x, x')) \tilde{\sigma}_{k,r,x,x'}^\diamond(E_r^\diamond(\varepsilon_r')^2).$$

If  $x$  and  $x'$  are both in  $E_r^\diamond(\tilde{\varepsilon}_r)$  then  $\tau(x, x') = 1$ , and the claim follows. When  $x \notin E_r^\diamond(\tilde{\varepsilon}_r)$  we get from (13.23) that  $\nu_{k,r,x}^{(n)}(E_r^\diamond(\varepsilon'_r)) = 0$ . Then

$$\tilde{\sigma}_{k,r,x,x'}^\diamond(E_r^\diamond(\varepsilon'_r)^2) \leq \sigma_{k,r,x}^\diamond(E_r^\diamond(\varepsilon'_r)) = \nu_{k,r,x}^{(n)}(E_r^\diamond(\varepsilon'_r)) = 0.$$

The case when  $x' \notin E_r^\diamond(\tilde{\varepsilon}_r)$  is analogous. Thus  $\hat{\sigma}_{k,r,x,x'}^\diamond(E_r^\diamond(\varepsilon'_r)^2) = 0$  also in case (b).

By part (ii) of Proposition 15.1, it follows that  $\tilde{\mathcal{Q}}_r \Psi_{k,r}(u, u') \leq \log \Omega_r + K_r$  for  $\hat{\sigma}_{k,r,x,x'}^\diamond$ -almost every  $(u, u') \in E_r^\diamond(\varepsilon_r)^2$ . Integrating with respect to  $\hat{\sigma}_{k,r,x,x'}^\diamond$  we get that  $\hat{\mathcal{T}}_{k,r}^\mathcal{Q} \Psi_{k,r}(x, x') \leq \log \Omega_r + K_r$  as claimed.  $\square$

**Proposition 16.4.** *There exist  $\kappa_r''' = \kappa_r'''(\nu_{\mathcal{L}}) > 0$  and  $C_r''' = C_r'''(\nu_{\mathcal{L}}) > 0$  such that given any  $\delta > 0$  and  $n \geq N_r$  the following holds for every  $k \geq \check{k}_r$ :*

(i) For any  $(x, x') \in E_r^\diamond(\varepsilon_r)^2 \setminus \text{Diag}_r$ ,

$$\hat{\mathcal{T}}_{k,r}^\mathcal{Q} \Psi_{k,r}(x, x') \leq \Psi_{k,r}(x, x') + C_r''' n.$$

(ii) For any  $(x, x') \in E_r^\diamond(\varepsilon_r)^2 \setminus \text{Diag}_r$  with  $\text{VA}_r(x, x') \geq \omega_{k,r}$ ,

$$\hat{\mathcal{T}}_{k,r}^\mathcal{Q} \Psi_{k,r}(x, x') \leq \Psi_{k,r}(x, x') + C_r'''(1 + \delta n).$$

(iii) For any  $(x, x') \in E_r^\diamond(\varepsilon_r'') \setminus \text{Diag}_r$  with  $\text{VA}_r(x, x') \geq \omega_{k,r}$ ,

$$\hat{\mathcal{T}}_{k,r}^\mathcal{Q} \Psi_{k,r}(x, x') \leq \Psi_{k,r}(x, x') - (\kappa_r''' - C_r''' \delta)n.$$

*Proof.* Take  $\kappa_r''' = \kappa_r''$  and  $C_r''' = C_r'' + K_r$ , and let  $k \geq \check{k}_r$ . We split the argument into four cases (check Figure 8).

First, suppose that both  $x$  and  $x'$  are in the  $\nu_k^{(n)}$ -core of  $E_r^\diamond(\varepsilon_r)$ , and at least one of them is in  $E_r^\diamond(\varepsilon_r'')$ . This is necessarily the case in the setting of (iii). In particular  $(x, x')$  is outside the cut-off region, which means that  $\Psi_{k,r}(x, x') > \log \Omega_r$ , and there is no recoupling either:

$$\hat{\sigma}_{k,r,x,x'}^\diamond = \nu_{k,r,x,x'}^{(n)}, \quad \hat{\mathcal{T}}_{k,r}^\mathcal{Q} \Psi_{k,r}(x, x') = \int_G \tilde{\mathcal{Q}}_r \Psi_{k,r}(gx, gx') d\nu_k^{(n)}(g).$$

Hence the claims in (i), (ii) and (iii) are contained in Corollary 15.5.

Now suppose that both  $x$  and  $x'$  are in the  $\nu_k^{(n)}$ -core of  $E_r^\diamond(\varepsilon_r)$  but outside  $E_r^\diamond(\varepsilon_r'')$ , and at least one of them is in  $E_r^\diamond(2\varepsilon_r'')$ . It is still true that  $(x, x')$  is outside the cut-off region, and so  $\Psi_{k,r}(x, x') > \log \Omega_r$ . Thus the estimates in Corollary 15.5 remain valid for

$$(16.13) \quad \int_{E_r^\diamond(\varepsilon_r)^2} \tilde{\mathcal{Q}}_r \Psi_{k,r} d\tilde{\sigma}_{k,r,x,x}^\diamond = \int_G \tilde{\mathcal{Q}}_r \Psi_{k,r}(gx, gx') d\nu_k^{(n)}(g).$$

By (16.9), the measure  $\zeta_{k,r,x,x'}$  vanishes on  $E_r^\diamond(\varepsilon_r'')^2$ . So, part (ii) of Proposition 15.1 gives that

$$(16.14) \quad \int_{E_r^\diamond(\varepsilon_r)^2} \tilde{\mathcal{Q}}_r \Psi_{k,r} d\zeta_{k,r,x,x} \leq \log \Omega_r + K_r \leq \Psi_{k,r}(x, x') + K_r.$$

By the definition (16.10),  $\hat{\mathcal{T}}_{k,r}^\mathcal{Q} \Psi_{k,r}(x, x')$  is a convex combination of the integrals in (16.13) and (16.14). Thus the claims (i) and (ii) follow in this case.

Next suppose that both  $x$  and  $x'$  are in the  $\nu_k^{(n)}$ -core of  $E_r^\diamond(\varepsilon_r)$  but outside  $E_r^\diamond(2\varepsilon_r'')$ . This corresponds to case (b) of Lemma 16.3: claims (i) and (ii) are contained in the conclusion of that lemma. Finally, suppose that at least one of

the points  $x$  and  $x'$  is in the  $\nu_k^{(n)}$ -border of  $E_r^\diamond(\varepsilon_r)$ . This is precisely the situation in case (a) of Lemma 16.3, and so claims (i) and (ii) are again contained in the conclusion of that lemma.  $\square$

**16.3. Contradicting**  $\dim E = r$ . Now we are going to apply Lemma 11.3 with  $X = E_r^\diamond(\varepsilon_r)^2$ ,  $\mathcal{T} = \widehat{\mathcal{T}}_{k,r}^\mathcal{Q}$ ,  $\psi = \Psi_{k,r}$ ,  $\hat{\eta} = \hat{\eta}_{k,r}$ ,

$$(16.15) \quad A_k = \{(x, x') \in E_r^\diamond(\varepsilon_r)^2 : \text{VA}_r(x, x') > \omega_{k,r}, \text{VA}_r(x', x) > \omega_{k,r}, \\ d(F_r, E) \leq \varepsilon_r''', \text{ and } d(F_r', E) \leq \varepsilon_r'''\},$$

$$(16.16) \quad B'_k = \{(x, x') \in E_r^\diamond(\varepsilon_r)^2 : \text{VA}_r(x, x') > \omega_{k,r}, \text{VA}_r(x', x) > \omega_{k,r}, \\ \text{and } d(F_r, E) > \varepsilon_r''' \text{ or } d(F_r', E) > \varepsilon_r'''\},$$

$$(16.17) \quad B_k^1 = \{(x, x') \in E_r^\diamond(\varepsilon_r)^2 : \text{VA}_r(x, x') \leq \omega_{k,r}\},$$

$$(16.18) \quad B_k^2 = \{(x, x') \in E_r^\diamond(\varepsilon_r)^2 : \text{VA}_r(x', x) \leq \omega_{k,r}\}.$$

The sets  $A = A_k$ ,  $B' = B'_k$ , and  $B'' = B_k^1 \cup B_k^2$  are pairwise disjoint, and their union is the whole  $E_r^\diamond(\varepsilon_r)^2$ . Moreover, (7.2) implies that  $B'' = \emptyset$  when  $\dim E = r$ .

Proposition 16.4 shows that, as long as  $k$  is sufficiently large, the hypotheses of Lemma 11.3 are satisfied for these choices, with

$$\kappa_A = (\kappa_r''' - C_r''' \delta)n, \quad \kappa'_B = C_r'''(1 + \delta)n, \quad \text{and } \kappa''_B = C_r'''n.$$

Take  $\delta > 0$  to be sufficiently small, depending on  $\nu_\infty$ , and  $n \in \mathbb{N}$  to be sufficiently large, depending on  $\nu_\infty$  and  $\delta$ , that

$$(16.19) \quad \kappa_A > 9\kappa'_B.$$

Using Proposition 6.6, we find a self-coupling  $\hat{\eta}_0^\diamond$  of  $\eta_{k,r}^\diamond$  vanishing on a neighborhood of  $\text{Diag}_r \subset E_r^\diamond(\varepsilon_r)^2$ . Observe that  $\hat{\eta}_0^\diamond$  projects to  $\eta_{k,r}$  under  $f \circ \pi_i^\diamond$  for any  $i = 1, 2$ . Then the same is true about every  $\widehat{\mathcal{T}}_{k,r}^\mathcal{Q}$ -iterate of  $\hat{\eta}_0^\diamond$ , by (16.12). Starting from  $\hat{\eta}_0^\diamond$  and arguing as in the proof of Proposition 6.23, we find a sequence  $(\hat{\eta}_{k,r,j}^\mathcal{Q})_j$  of probability measures on  $E_r^\diamond(\varepsilon_r)^2$  projecting to  $\eta_{k,r}$  under  $f \circ \pi_i^\diamond$  for any  $i = 1, 2$ , converging to a  $\widehat{\mathcal{T}}_{k,r}^\mathcal{Q}$ -invariant measure  $\hat{\eta}_{k,r}^\mathcal{Q}$ , and satisfying  $\int_{E_r^\diamond(\varepsilon_r)^2} \Psi_{k,r} d\hat{\eta}_{k,r,j}^\mathcal{Q} < \infty$  and

$$\int_{E_r^\diamond(\varepsilon_r)^2} \widehat{\mathcal{T}}_{k,r}^\mathcal{Q} \Psi_{k,r}(x, x') d\hat{\eta}_{k,r,j}^\mathcal{Q}(x, x') \geq \int_{E_r^\diamond(\varepsilon_r)^2} \Psi_{k,r}(x, x') d\hat{\eta}_{k,r,j}^\mathcal{Q}(x, x')$$

for every  $j$ . Applying Lemma 11.3 with  $\mathcal{T} = \widehat{\mathcal{T}}_{k,r}^\mathcal{Q}$  and  $\hat{\eta} = \hat{\eta}_{k,r,j}^\mathcal{Q}$ , we get that

$$(16.20) \quad \hat{\eta}_{k,r,j}^\mathcal{Q}(B'') \geq \frac{\kappa_A \hat{\eta}_{k,r,j}^\mathcal{Q}(E_r^\diamond(\varepsilon_r)^2) - (\kappa_A + \kappa'_B) \hat{\eta}_{k,r,j}^\mathcal{Q}(B')}{\kappa_A + \kappa''_B}$$

for every  $j$ . Passing to the limit as  $j \rightarrow \infty$ , we conclude that

$$(16.21) \quad \hat{\eta}_{k,r}^\mathcal{Q}(B'') \geq \frac{\kappa_A \hat{\eta}_{k,r}^\mathcal{Q}(E_r^\diamond(\varepsilon_r)^2) - (\kappa_A + \kappa'_B) \hat{\eta}_{k,r}^\mathcal{Q}(B')}{\kappa_A + \kappa''_B}.$$

By definition,  $B'$  is contained in the union of the pre-images  $(f \circ \pi_i^\diamond)^{-1}(E_r(\varepsilon_r, \varepsilon_r'''))$ ,  $i = 1, 2$ . Thus, using (16.3),

$$(16.22) \quad \hat{\eta}_{k,r}^\mathcal{Q}(B') \leq 2\eta_{k,r}(E_r(\varepsilon_r, \varepsilon_r''')) < \frac{4}{10}\eta_{k,r}(E_r(\varepsilon_r)).$$



It is clear that  $\hat{\eta}_{k,r}^{\mathcal{Q}}(E_r^{\circ}(\varepsilon_r)^2) = \eta_{k,r}(E_r(\varepsilon_r))$ . Substituting these relations in (16.21) and using (16.19), we find that

$$(16.23) \quad \hat{\eta}_{k,r}^{\mathcal{Q}}(B'') \geq \frac{\kappa_A - \frac{4}{10}(\kappa_A + \kappa'_B)}{\kappa_A + \kappa''_B} \eta_{k,r}(E_r(\varepsilon_r)) \geq \frac{5\kappa'_B}{\kappa_A + \kappa''_B} \eta_{k,r}(E_r(\varepsilon_r)) > 0.$$

When  $\dim E = r$  this is a contradiction, because  $B''$  is empty in that case. Thus  $\dim E \geq r + 1$ .

**16.4. Completing step  $r$ .** By (16.23), there exists  $i = 1, 2$  such that

$$(16.24) \quad \hat{\eta}_{k,r}^{\mathcal{Q}}(B_k^i) \geq \frac{2\kappa'_B}{\kappa_A + \kappa''_B} \eta_{k,r}(E_r(\varepsilon_r)).$$

It is no restriction to assume that  $i = 1$ , as the other case can be deduced just by exchanging the roles of  $x$  and  $x'$ .

Consider the map

$$\Sigma : E_r^{\circ}(\varepsilon_r)^2 \setminus \text{Diag}_r \rightarrow \text{Gr}(r+1, d), \quad \Sigma(x, x') = F'_1 + F_r$$

and the compact topological spaces

$$\begin{aligned} \mathcal{Y}_r &= \{(x, x', y) \in \mathcal{F}(r, d)^2 \times \text{Gr}(r+1, d) : F'_1 \subset y \text{ and } F_r \subset y\} \\ \mathcal{Y}_r(\varepsilon) &= \{(x, x', y) \in \mathcal{Y}_r : x, x' \in E_r^{\circ}(\varepsilon)\} \text{ for } \varepsilon > 0, \end{aligned}$$

together with the canonical projections

$$\begin{aligned} p_1 : \mathcal{Y}_r &\rightarrow \mathcal{F}(r, d)^2, & (x, x', y) &\mapsto (x, x') \\ p_2 : \mathcal{Y}_r &\rightarrow \text{Gr}(r+1, d), & (x, x', y) &\mapsto y. \end{aligned}$$

For  $(x, x', y) \in \mathcal{Y}_r$  and  $n \in \mathbb{N}$ , denote by  $\nu_{k,r,x,x',y}^{(n)}$  the image of  $\nu_k^{(n)}$  under the diagonal action

$$G \rightarrow \mathcal{Y}_r, \quad (g \mapsto (gx, gx', gy)).$$

Clearly, each  $\nu_{k,r,x,x',y}^{(n)}$  is a lift of  $\nu_{k,r,x,x'}^{(n)}$  relative to  $p_1 : \mathcal{Y}_r \rightarrow \mathcal{F}(r, d)^2$ . The complement of  $\text{Diag}_r$  in  $\mathcal{F}(r, d)^2$  embeds in  $\mathcal{Y}_r$  through

$$(x, x') \mapsto (x, x', F'_1 + F_r).$$

In particular, every measure  $\xi$  on  $\mathcal{F}(r, d)^2$  that vanishes on  $\text{Diag}_r$  has a (unique) lift  $\tilde{\xi}$  to  $\mathcal{Y}_r$ .

From the relations (11.3) and (16.10), we see that

$$(16.25) \quad \hat{\sigma}_{k,r,x,x'}^{\circ} = (1 - \tilde{\omega}(x, x')) \nu_{k,r,x,x'}^{(n)} + \tilde{\omega}(x, x') \hat{\theta}_{k,r,x,x'}$$

where  $\tilde{\omega} : E_r^{\circ}(\varepsilon_r)^2 \rightarrow [0, 1]$  is a continuous function that vanishes identically on  $E_r^{\circ}(\varepsilon_r')^2$ , and each  $\hat{\theta}_{k,r,x,x'}$  is a coupling of  $\sigma_{k,1,x}^{\circ}$  and  $\sigma_{k,1,x'}^{\circ}$  vanishing on a uniform neighborhood of  $\text{Diag}_r$ . In view of the previous remarks, it follows that the  $\hat{\sigma}_{k,r,x,x'}^{\circ}$  lift to probability measures

$$(16.26) \quad \check{\sigma}_{k,r,x,x',y} = (1 - \tilde{\omega}(x, x')) \nu_{k,r,x,x',y}^{(n)} + \tilde{\omega}(x, x') \check{\theta}_{k,r,x,x',y}$$

on  $\mathcal{Y}_r(\varepsilon_r)$ , where  $\check{\theta}_{k,r,x,x',y}$  is the unique lift of  $\hat{\theta}_{k,r,x,x',y}$ . This lift is continuous: it is clear that  $\check{\nu}_{k,r,x,x',y}^{(n)}$  varies continuously and, by uniqueness, so does  $\check{\theta}_{k,r,x,x',y}$ .

We claim that the spreading out measures  $\check{q}_{r,x,x'}$  in (15.6) also lift continuously to measures  $\check{q}_{r,x,x',y}$  on  $\mathcal{Y}_r(\varepsilon_r)$ . Indeed, it is clear that  $\{\delta_{(x,x',y)} : (x, x', y) \in \mathcal{Y}_r(\varepsilon_r)\}$

is a continuous lift of  $\{\delta_{(x,x')} : (x, x') \in E_r^\circ(\varepsilon_r)^2\}$ , and so it suffices to show that the family

$$\{\mu_{F_r}^* \times \mu_{F'_r}^* : (x, x') \in E_r^\circ(\varepsilon_r, \varepsilon'_r/2) \times E_r^\circ(3\varepsilon_r'') \cup E_r^\circ(3\varepsilon_r'') \times E_r^\circ(\varepsilon_r, \varepsilon'_r/2)\}$$

lifts uniquely to  $\mathcal{Y}_r(\varepsilon_r)$ . The latter is a direct consequence of the fact that

$$(16.27) \quad (\mu_{F_r}^* \times \mu_{F'_r}^*) (\text{Diag}_r) = 0$$

for any  $(F_r, F'_r) \in E_r(\varepsilon_r, \varepsilon'_r/2) \times E_r(3\varepsilon_r'') \cup E_r(3\varepsilon_r'') \times E_r(\varepsilon_r, \varepsilon'_r/2)$ . To prove (16.27), let us write  $u = (G_1, \dots, G_r)$  and  $u' = (G'_1, \dots, G'_r)$ . By definition,

$$(16.28) \quad (\mu_{F_r}^* \times \mu_{F'_r}^*) (\text{Diag}_r) \leq \mu_{F_r}^* (\{u' \in \mathcal{F}(F'_r) : G'_1 \subset F_r\}) \\ + \mu_{F'_r}^* (\{u \in \mathcal{F}(F_r) : G_1 \subset F'_r\})$$

The key point is that in this setting we always have  $F_r \neq F'_r$ . Thus the set of  $u' \in \mathcal{F}(F'_r)$  such that  $G'_1 \subset F_r$  is a subvariety of  $\mathcal{F}(F'_r)$  of strictly smaller dimension, and so it has zero  $\mu_{F_r}^*$ -measure. Thus the first term on the right-hand side of (16.28) vanishes identically, and then so does the second term, by symmetry. This proves the claim.

These observations ensure that the Markov operator  $\widehat{\mathcal{T}}_{k,r}^\mathcal{Q} = \widehat{\mathcal{T}}_{k,r} \circ \widetilde{\mathcal{Q}}_r$  admits a continuous lift

$$\widetilde{\mathcal{T}}_{k,r}^\mathcal{Q} : \mathcal{B}(\mathcal{Y}_r(\varepsilon_r)) \rightarrow \mathcal{B}(\mathcal{Y}_r(\varepsilon_r)), \quad \widetilde{\mathcal{T}}_{k,r}^\mathcal{Q} \Phi(x, x', y) = \int_{\mathcal{Y}_r(\varepsilon_r)} \Phi d\check{\sigma}_{k,r,x,x',y}^\mathcal{Q},$$

given by

$$\check{\sigma}_{k,r,x,x',y}^\mathcal{Q} = \int_{\mathcal{Y}_r(\varepsilon_r)} \check{q}_{r,u,u',v} d\check{\sigma}_{k,r,x,x',y}(u, u', v).$$

Since the measure  $\hat{\eta}_0^\circ$  was taken to vanish on a neighborhood of  $\text{Diag}_r \subset E_r^\circ(\varepsilon_r)^2$ , it also admits a (unique) lift  $\check{\eta}_0^\mathcal{Q}$  to  $\mathcal{Y}_r(\varepsilon_r)$ . Applying the construction in Proposition 6.23 simultaneously to the operators  $\widehat{\mathcal{T}}_{k,r}^\mathcal{Q}$  and  $\widetilde{\mathcal{T}}_{k,r}^\mathcal{Q}$ , starting from  $\hat{\eta}_0^\circ$  and  $\check{\eta}_0^\mathcal{Q}$  respectively, we find a sequence  $(\check{\eta}_{k,r,j}^\mathcal{Q})_j$  of probability measures on  $\mathcal{Y}_r(\varepsilon_r)$  converging to a  $\widetilde{\mathcal{T}}_{k,r}^\mathcal{Q}$ -invariant probability measure  $\check{\eta}_{k,r}^\mathcal{Q}$  and such that (up to restricting to a subsequence) each  $\check{\eta}_{k,r,j}^\mathcal{Q}$  projects to  $\hat{\eta}_{k,r,j}^\mathcal{Q}$  under  $p_1$ . Observe that  $\check{\eta}_{k,r}^\mathcal{Q}$  and each  $\check{\eta}_{k,r,j}^\mathcal{Q}$  project to  $\eta_{k,r}$  under  $f \circ \pi_i^\circ \circ p_1$  for  $i = 1, 2$ .

Next, define  $\eta_{k,r+1} = p_{2*} \check{\eta}_{k,r}^\mathcal{Q}$  and let  $\{d\check{\eta}_{k,r,v} : v \in p_2 \mathcal{Y}_r(\varepsilon_r)\}$  be a disintegration of  $\check{\eta}_{k,r}^\mathcal{Q}$  with respect to the partition  $\{p_2^{-1}(v) : v \in p_2 \mathcal{Y}_r(\varepsilon_r)\}$ . Then define

$$\mathcal{T}_{k,r+1} : \mathcal{B}(p_2 \mathcal{Y}_r(\varepsilon_r)) \rightarrow \mathcal{B}(p_2 \mathcal{Y}_r(\varepsilon_r)), \\ \mathcal{T}_{k,r+1} \Phi(y) = \int_{p_2^{-1}(y)} \widetilde{\mathcal{T}}_{k,r}^\mathcal{Q} (\Phi \circ p_2)(x, x', y) d\check{\eta}_{k,r,y}(x, x').$$

Equivalently,  $\mathcal{T}_{k,r+1} \Phi(y) = \int_{p_2 \mathcal{Y}_r(\varepsilon_r)} \Phi d\sigma_{k,r+1,y}$  with

$$(16.29) \quad \sigma_{k,r+1,y} = \int_{p_2^{-1}(y)} p_{2*} \check{\sigma}_{k,r,x,x',y}^\mathcal{Q} d\check{\eta}_{k,r,y}(x, x').$$

Let  $\mathcal{B}_k = p_2 p_1^{-1}(B_k^1) = \{F'_1 + F_r : (x, x') \in B_k^1\}$ , where  $B_k^1$  is as in the previous section. Define also  $\eta_{k,r+1,j} = p_{2*} \check{\eta}_{k,r,j}^\mathcal{Q}$  for  $j \in \mathbb{N}$ . Then

$$\eta_{k,r+1,j}(\mathcal{B}_k) \geq \check{\eta}_{k,r,j}^\mathcal{Q}(p_1^{-1}(B_k^1)) = \hat{\eta}_{k,r,j}^\mathcal{Q}(B_k^1).$$

Passing to the limit as  $j \rightarrow \infty$ , and arguing as in (16.20)–(16.23), we find from (16.24) that

$$\eta_{k,r+1}(\mathcal{B}_k) \geq \frac{5\kappa'_B}{\kappa_A + \kappa''_B} \eta_{k,r}(E_r(\varepsilon_r)).$$

Now, since  $\omega_{k,r} \rightarrow 0$ , the definition (16.17) implies that  $\mathcal{B}_k$  converges to  $E_r$  as  $k \rightarrow \infty$ . Thus, any accumulation point  $\eta_{\infty,r+1}$  of  $\eta_{k,r+1}$  must satisfy

$$(16.30) \quad \eta_{\infty,r+1}(E_{r+1}) \geq \frac{5\kappa'_B}{\kappa_A + \kappa''_B} \eta_{\infty,r}(E_r(\varepsilon_r)) \geq \frac{5\kappa'_B}{\kappa_A + \kappa''_B} \eta_{\infty,r}(E_r) > 0.$$

Take  $n_{r+1} = n$  and  $\varepsilon_{r+1} = \varepsilon_r''$ . Let  $\nu_{k,r+1,y}^{(n_{r+1})}$  denote the push-forward of  $\nu_k^{(n_{r+1})}$  under the map  $G \rightarrow \text{Gr}(r+1, d)$ ,  $g \mapsto gy$ .

**Lemma 16.5.**

- (i)  $\sigma_{k,r+1,y} = \nu_{k,r+1,y}^{(n_{r+1})}$  for every  $y \in E_{r+1}(\varepsilon_{r+1})$ .
- (ii)  $\sigma_{k,r+1,y} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_{r+1}(\varepsilon_{r+1}) \right) = 0$  for every  $y \notin E_{r+1}(\varepsilon_{r+1})$
- (iii) The measure  $\eta_{k,r+1}$  is  $\mathcal{T}_{k,r+1}$ -invariant.

*Proof.* It is clear that  $\nu_{k,r+1,y}^{(n_{r+1})}$  coincides with the push-forward of  $\nu_{k,r,x,x',y}^{(n_{r+1})}$  under the projection  $p_2$ . Thus (16.26) gives that

$$p_{2*} \check{\sigma}_{k,r,x,x',y} = (1 - \check{\omega}(x, x')) \nu_{k,r+1,y}^{(n_{r+1})} + \check{\omega}(x, x') p_{2*} \check{\theta}_{k,r,x,x',y},$$

and so,

$$\begin{aligned} \check{\sigma}_{k,r+1,y} &= \left( 1 - \int_{p_2^{-1}(y)} \check{\omega}(x, x') d\check{\eta}_{k,r,y}(x, x') \right) \nu_{k,r+1,y}^{(n_{r+1})} \\ &\quad + \int_{p_2^{-1}(y)} \check{\omega}(x, x') p_{2*} \check{\theta}_{k,r,x,x',y} d\check{\eta}_{k,r,y}(x, x'), \end{aligned}$$

If  $y \in E_{r+1}(\varepsilon_{r+1})$  then both  $x$  and  $x'$  are necessarily in  $E_r^\diamond(\varepsilon_{r+1})$ , by (7.1), in which case  $\check{\omega}(x, x') = 0$ . Then  $\sigma_{k,r+1,y} = \nu_{k,r+1,y}^{(n_{r+1})}$ , as claimed in (i).

In view of (16.29), to prove part (ii) it suffices to show that if  $y \notin E_{r+1}(\varepsilon_{r+1})$  then

$$(16.31) \quad p_{2*} \check{\sigma}_{k,r,x,x',y} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_{r+1}(\varepsilon_{r+1}) \right) = 0$$

for any  $x, x' \subset y$ . If  $x$  and  $x'$  are both in  $E_r^\diamond(\varepsilon_{r+1})$  then

$$p_{2*} \check{\sigma}_{k,r,x,x',y} = p_{2*} \nu_{k,r,x,x',y}^{(n_{r+1})} = \nu_{k,r+1,y}^{(n_{r+1})}$$

and then the claim follows from Remark 6.26. From now on, we assume that one of the points,  $x$  say, is not in  $E_r^\diamond(\varepsilon_{r+1})$ . It follows from the definitions that

$$p_2^{-1} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_{r+1}(\varepsilon_{r+1}) \right) \subset \mathcal{X}_{\nu_k^{(n)}}^\# E_r^\diamond(\varepsilon_{r+1})^2 \times \mathcal{X}_{\nu_k^{(n)}}^\# E_{r+1}(\varepsilon_{r+1}),$$

and so

$$\begin{aligned}
(16.32) \quad p_{2*} \check{\sigma}_{k,r,x,x',y} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_{r+1}(\varepsilon_{r+1}) \right) \\
\leq \check{\sigma}_{k,r,x,x',y} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_r^\diamond(\varepsilon_{r+1})^2 \times \mathcal{X}_{\nu_k^{(n)}}^\# E_{r+1}(\varepsilon_{r+1}) \right) \\
\leq \hat{\sigma}_{k,r,x,x'}^\diamond \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_r^\diamond(\varepsilon_{r+1})^2 \right) \\
\leq \sigma_{k,r,x}^\diamond \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_r^\diamond(\varepsilon_{r+1}) \right).
\end{aligned}$$

If  $x$  is in the  $\nu_k^{(n)}$ -border of  $E_r^\diamond(\varepsilon_r)$  then

$$\sigma_{k,r,x}^\diamond \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_r^\diamond(\varepsilon_{r+1}) \right) \leq \sigma_{k,r,x}^\diamond \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_r^\diamond(\varepsilon_r) \right) = 0,$$

because the operator  $\mathcal{T}_{k,1}$  is adapted to  $(\nu_k^{(n)}, E_r^\diamond(\varepsilon_r))$ . If  $x$  is in the  $\nu_k^{(n)}$ -core of  $E_r^\diamond(\varepsilon_r)$  then Remark 6.26 gives that

$$\sigma_{k,r,x}^\diamond \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_r^\diamond(\varepsilon_{r+1}) \right) = \nu_{k,r,x}^{(n)} \left( \mathcal{X}_{\nu_k^{(n)}}^\# E_r^\diamond(\varepsilon_{r+1}) \right) = 0.$$

Thus the right-hand side of (16.32) vanishes in either case. That completes the proof of (16.31) and of part (ii) of the lemma.

Finally, by definition,

$$\begin{aligned}
\int_{p_2 \mathcal{Y}_r(\varepsilon_r)} (\mathcal{T}_{k,r+1} \Phi) d\eta_{k,r+1} \\
= \int_{p_2 \mathcal{Y}_r(\varepsilon_r)} \int_{p_2^{-1}(y)} \check{\mathcal{T}}_{k,r}(\Phi \circ p_2)(x, x', y) d\check{\eta}_{k,r,y}(x, x') d\eta_{k,r+1}(y) \\
= \int_{\mathcal{Y}_r(\varepsilon_r)} \check{\mathcal{T}}_{k,r}(\Phi \circ p_2)(x, x', y) d\check{\eta}_{k,r}(x, x', y)
\end{aligned}$$

for any  $\Phi \in \mathcal{B}(p_2 \mathcal{Y}_r(\varepsilon_r))$ . Since  $\check{\eta}_{k,r}$  is  $\check{\mathcal{T}}_{k,r}$ -invariant, this gives

$$\begin{aligned}
\int_{p_2 \mathcal{Y}_r(\varepsilon_r)} (\mathcal{T}_{k,r+1} \Phi)(y) d\eta_{k,r+1}(y) &= \int_{\mathcal{Y}_r(\varepsilon_r)} (\Phi \circ p_2)(x, x', y) d\check{\eta}_{k,r}(x, x', y) \\
&= \int_{p_2 \mathcal{Y}_r(\varepsilon_r)} \Phi d\hat{\eta}_{k,r+1},
\end{aligned}$$

which proves claim (iii).  $\square$

Since the  $\check{\sigma}_{k,r,x,x',y}$  are generic measures and the projection  $p_2$  is algebraic, it follows readily from (16.29) and Remark 5.1 that every  $\sigma_{k,r+1,y}$  is a generic measure. Then, conclusions (i) and (ii) in Lemma 16.5 allow us to apply Propositions 6.15 and 6.25 with  $E_{r+1}(\varepsilon_{r+1})$  and  $p_2 \mathcal{Y}_r(\varepsilon_r)$  in the roles of  $X$  and  $U$ , respectively. Thus we get a continuous Markov operator adapted to  $(\nu_k^{(n_{r+1})}, E_{r+1}(\varepsilon_{r+1}))$  and which leaves the restriction of  $\eta_{k,r+1} \mid E_{r+1}(\varepsilon_{r+1})$  invariant. Replace  $\mathcal{T}_{k,r+1}$  and  $\eta_{k,r+1}$  with these new Markov operator and invariant measure, respectively. This finishes step  $r$  of the induction.

The proof of Theorem 7.1 is now complete.

## REFERENCES

- [1] J. Athreya. Quantitative recurrence and large deviations for Teichmüller geodesic flow. *Geom. Dedicata*, 119:121–140, 2006.
- [2] A. Avila and G. Forni. Weak mixing for interval exchange transformations and translation flows. *Ann. of Math.*, 165:637–664, 2007.
- [3] A. Avila and S. Gouëzel. Small eigenvalues of the Laplacian for algebraic measures in moduli space, and mixing properties of the Teichmüller flow. *Ann. of Math.*, 178:385–442, 2013.
- [4] A. Avila and S. Jitomirskaya. The Ten Martini Problem. *Ann. of Math. (2)*, 170:303–342, 2009.
- [5] A. Avila, J. Santamaria, and M. Viana. Holonomy invariance: rough regularity and applications to Lyapunov exponents. *Astérisque*, 358:13–74, 2013.
- [6] A. Avila and M. Viana. Simplicity of Lyapunov spectra: proof of the Zorich-Kontsevich conjecture. *Acta Math.*, 198:1–56, 2007.
- [7] A. Avila and M. Viana. Extremal Lyapunov exponents: an invariance principle and applications. *Invent. Math.*, 181:115–189, 2010.
- [8] A. Avila, M. Viana, and A. Wilkinson. Absolute continuity, Lyapunov exponents and rigidity I: geodesic flows. *J. Eur. Math. Soc. (JEMS)*, 17:1435–1462, 2015.
- [9] A. Avila, M. Viana, and A. Wilkinson. Absolute continuity, Lyapunov exponents, and rigidity II: systems with compact center leaves. *Ergodic Theory Dynam. Systems*, 42:437–490, 2022.
- [10] Artur Avila. Global theory of one-frequency Schrödinger operators. *Acta Math.*, 215:1–54, 2015.
- [11] K. Azuma. Weighted sums of certain dependent random variables. *Tôhoku Math. J.*, 19:357–367, 1967.
- [12] L. Backes, A. Brown, and C. Butler. Continuity of Lyapunov exponents for cocycles with invariant holonomies. *Journal of Modern Dynamics*, 12:223–260, 2018.
- [13] L. Backes and M. Poletti. Continuity of Lyapunov exponents is equivalent to continuity of Oseledec subspaces. *Stoch. Dyn.*, 17:1750047, 18, 2017.
- [14] Y. Benoist and J.-F. Quint. Random walks on finite volume homogeneous spaces. *Invent. Math.*, 187:37–59, 2012.
- [15] J. Bochi. Genericity of zero Lyapunov exponents. *Ergod. Th. & Dynam. Sys.*, 22:1667–1696, 2002.
- [16] J. Bochi.  $C^1$ -generic symplectic diffeomorphisms: partial hyperbolicity and zero centre Lyapunov exponents. *J. Inst. Math. Jussieu*, 8:49–93, 2009.
- [17] J. Bochi and M. Viana. The Lyapunov exponents of generic volume-preserving and symplectic maps. *Ann. of Math.*, 161:1423–1485, 2005.
- [18] C. Bocker and M. Viana. Continuity of Lyapunov exponents for random two-dimensional matrices. *Ergodic Theory Dynam. Systems*, 37:1413–1442, 2017.
- [19] C. Bonatti, A. Eskin, and A. Wilkinson. Projective cocycles over  $SL(2, \mathbb{R})$  actions: measures invariant under the upper triangular group. *Astérisque*, 415:157–180, 2020.
- [20] C. Bonatti, X. Gómez-Mont, and M. Viana. Généricité d’exposants de Lyapunov non-nuls pour des produits déterministes de matrices. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 20:579–624, 2003.
- [21] J. Bourgain. Positivity and continuity of the Lyapunov exponent for shifts on  $\mathbb{T}^d$  with arbitrary frequency vector and real analytic potential. *J. Anal. Math.*, 96:313–355, 2005.
- [22] J. Bourgain. An application of group expansion to the Anderson-Bernoulli model. *GAFA*, 24:49–62, 2014.
- [23] J. Bourgain and S. Jitomirskaya. Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential. *J. Statist. Phys.*, 108:1203–1218, 2002.
- [24] P. Buterus, F. Götze, T. Hille, and G. Margulis. Distribution of values of quadratic forms at integral points. *Invent. Math.*, 227:857–961, 2022.
- [25] C. Butler. Discontinuity of Lyapunov exponents near fiber bunched cocycles. *Ergodic Theory Dynam. Systems*, 38:523–539, 2018.
- [26] P. Duarte and S. Klein. *Lyapunov exponents of linear cocycles*, volume 3 of *Atlantis Studies in Dynamical Systems*. Atlantis Press, 2016. Continuity via large deviations.
- [27] P. Duarte and S. Klein. Continuity, positivity and simplicity of the Lyapunov exponents for quasi-periodic cocycles. *J. Eur. Math. Soc.*, 21:2051–2106, 2019.

- [28] P. Duarte and S. Klein. Large deviations for products of random two dimensional matrices. *Comm. Math. Physics*, 375:2191–2257, 2020.
- [29] N. Dunford and J. Schwarz. *Linear operators I: General theory*. Wiley & Sons, 1957.
- [30] M. Einsiedler and S. Kadyrov. Entropy and escape of mass for  $SL_3(\mathbb{Z}) \backslash SL_3(\mathbb{R})$ . *Israel J. Math.*, 190:253–288, 2012.
- [31] A. Eskin and G. Margulis. Recurrence properties of random walks on finite volume homogeneous manifolds. In *Random walks and geometry*, pages 431–444. Walter de Gruyter, 2004.
- [32] A. Eskin, G. Margulis, and S. Mozes. Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture. *Ann. of Math.*, 147:93–141, 1998.
- [33] A. Eskin, G. Margulis, and S. Mozes. Quadratic forms of signature  $(2, 2)$  and eigenvalue spacings on rectangular 2-tori. *Ann. of Math.*, 161:679–725, 2005.
- [34] A. Eskin and H. Masur. Asymptotic formulas on flat surfaces. *Ergodic Theory Dynam. Systems*, 21:443–478, 2001.
- [35] A. Eskin and M. Mirzakhani. Counting closed geodesics in moduli space. *J. Mod. Dyn.*, 5:71–105, 2011.
- [36] A. Eskin and M. Mirzakhani. Invariant and stationary measures for the  $SL(2, \mathbb{R})$  action on moduli space. *Publ. Math. IHES*, 127:95–234, 2018.
- [37] A. Eskin, M. Mirzakhani, and A. Mohammadi. Isolation, equidistribution and orbit closures for the  $SL(2, \mathbb{R})$  action on moduli space. *Ann. of Math.*, 182:673–721, 2015.
- [38] A. Eskin, S. Mozes. Margulis functions and their applications. Dynamics, geometry, number theory—the impact of Margulis on modern mathematics, 342–361, Univ. Chicago Press, Chicago, IL, [2022].
- [39] A. Fathi, M. Herman, and J.-C. Yoccoz. A proof of Pesin’s stable manifold theorem. In *Geometric dynamics (Rio de Janeiro 1981)*, volume 1007 of *Lect. Notes in Math.*, pages 177–215. Springer Verlag, 1983.
- [40] H. Furstenberg. Non-commuting random products. *Trans. Amer. Math. Soc.*, 108:377–428, 1963.
- [41] H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960.
- [42] H. Furstenberg and Yu. Kifer. Random matrix products and measures in projective spaces. *Israel J. Math.*, 10:12–32, 1983.
- [43] A. Gorodetski, V. Kleptsyn, G. Monakov. Hölder regularity of stationary measures. [arXiv:2209.12342 \[math.DS\]](https://arxiv.org/abs/2209.12342).
- [44] T. Gelander, T. A. Levit, G. A. Margulis. Effective discreteness radius of stabilizers for stationary actions. *Michigan Math. J.* 72 (2022), 389–438.
- [45] I. Ya. Gol’sheid and G. A. Margulis. Lyapunov indices of a product of random matrices. *Uspekhi Mat. Nauk.*, 44:13–60, 1989.
- [46] Y. Guivarc’h and A. Raugi. Products of random matrices : convergence theorems. *Contemp. Math.*, 50:31–54, 1986.
- [47] H. Hennion. Loi des grands nombres et perturbations pour des produits réductibles de matrices aléatoires indépendantes. *Z. Wahrsch. Verw. Gebiete*, 67:265–278, 1984.
- [48] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.
- [49] S. Kadyrov, D. Kleinbock, E. Lindenstrauss, and G. Margulis. Entropy in the cusp and singular systems of linear forms. *J. Anal. Math.*, 133:253–277, 2017.
- [50] Yu. Kifer. Perturbations of random matrix products. *Z. Wahrsch. Verw. Gebiete*, 61:83–95, 1982.
- [51] F. Ledrappier. Positivity of the exponent for stationary sequences of matrices. In *Lyapunov exponents (Bremen, 1984)*, volume 1186 of *Lect. Notes Math.*, pages 56–73. Springer-Verlag, 1986.
- [52] A. M. Lyapunov. *The general problem of the stability of motion*. Taylor & Francis Ltd., 1992. Translated from Edouard Davaux’s French translation (1907) of the 1892 Russian original and edited by A. T. Fuller. With an introduction and preface by Fuller, a biography of Lyapunov by V. I. Smirnov, and a bibliography of Lyapunov’s works compiled by J. F. Barrett. Lyapunov centenary issue. Reprint of *Internat. J. Control* 55 (1992), no. 3. With a foreword by Ian Stewart.
- [53] E. Malheiro and M. Viana. Lyapunov exponents of linear cocycles over Markov shifts. *Stoch. Dyn.*, 15:1550020, 27, 2015.

- [54] R. Mañé. Oseledec's theorem from the generic viewpoint. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983)*, pages 1269–1276. PWN, 1984.
- [55] S. Meyn, R. Tweedie. Markov chains and stochastic stability. Second edition. With a prologue by Peter W. Glynn. Cambridge University Press, Cambridge, 2009. xxviii+594 pp. ISBN: 978-0-521-73182-9
- [56] V. I. Oseledets. A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, 19:197–231, 1968.
- [57] É. Le Page. Régularité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes et applications. *Ann. Inst. H. Poincaré Probab. Statist.*, 25:109–142, 1989.
- [58] Y. Peres. Analytic dependence of Lyapunov exponents on transition probabilities. In *Lyapunov exponents (Oberwolfach, 1990)*, volume 1486 of *Lecture Notes in Math.*, pages 64–80. Springer-Verlag, 1991.
- [59] M. Poletti and M. Viana. Simple Lyapunov spectrum for certain linear cocycles over partially hyperbolic maps. *Nonlinearity*, 32:238–284, 2019.
- [60] C. Pugh and M. Shub. Ergodic attractors. *Trans. Amer. Math. Soc.*, 312:1–54, 1989.
- [61] D. Ruelle. Analyticity properties of the characteristic exponents of random matrix products. *Adv. in Math.*, 32:68–80, 1979.
- [62] B. Simon and M. Taylor. Harmonic analysis on  $SL(2, \mathbf{R})$  and smoothness of the density of states in the one-dimensional Anderson model. *Comm. Math. Phys.*, 101:1–19, 1985.
- [63] E. H. Y. Tall and M. Viana. Moduli of continuity for the Lyapunov exponents of random  $GL(2)$ -cocycles. *Trans. Amer. Math. Soc.*, 373:1343–1383, 2020.
- [64] M. Viana. Almost all cocycles over any hyperbolic system have nonvanishing Lyapunov exponents. *Ann. of Math.*, 167:643–680, 2008.
- [65] M. Viana. *Lectures on Lyapunov exponents*, volume 145 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2014.
- [66] M. Viana and J. Yang. Continuity of Lyapunov exponents in the  $C^0$  topology. *Israel J. Math.*, 229:461–485, 2019.
- [67] Y. Wang and J. You. Examples of discontinuity of Lyapunov exponent in smooth quasiperiodic cocycles. *Duke Math. J.*, 13:2363–2412, 2013.

IMPA, EST. D. CASTORINA 110, JARDIM BOTÂNICO, 22460-320 RIO DE JANEIRO, BRAZIL AND  
 INSTITUT FÜR MATHEMATIK, UNIVERSITÄT ZÜRICH, WINTERTHURERSTRASSE 190, 8057 ZÜRICH,  
 SWITZERLAND

*Email address:* `avila@impa.br`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, CHICAGO, IL 60637, USA

*Email address:* `eskin@math.uchicago.edu`

IMPA, EST. D. CASTORINA 110, JARDIM BOTÂNICO, 22460-320 RIO DE JANEIRO, BRAZIL

*Email address:* `viana@impa.br`